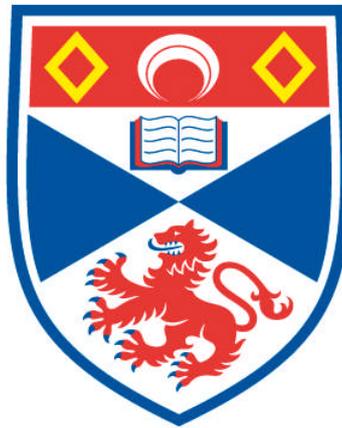


# **COUPLING OF THE SOLAR WIND, MAGNETOSPHERE AND IONOSPHERE BY MHD WAVES**

**Alexander J. B. Russell**

**A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews**



**2010**

**Full metadata for this item is available in  
Research@StAndrews:FullText  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

**Please use this identifier to cite or link to this item:**

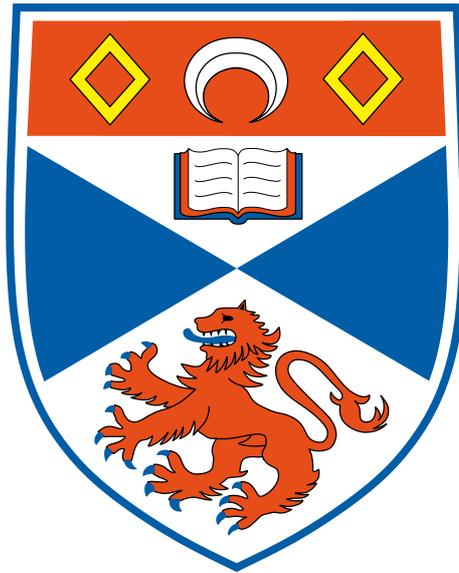
**<http://hdl.handle.net/10023/2571>**

**This item is protected by original copyright**

**This item is licensed under a  
Creative Commons License**

# **Coupling of the Solar Wind, Magnetosphere and Ionosphere by MHD Waves**

**Alexander J. B. Russell**



Thesis submitted for the degree of Doctor of Philosophy  
of the University of St Andrews

3 September 2010

---

# Abstract

---

The solar wind, magnetosphere and ionosphere are coupled by magnetohydrodynamic waves, and this gives rise to new and often unexpected behaviours that cannot be produced by a single, isolated part of the system. This thesis examines two broad instances of coupling: field-line resonance (FLR) which couples fast and Alfvén waves, and magnetosphere-ionosphere (MI-) coupling via Alfvén waves.

The first part of this thesis investigates field-line resonance for equilibria that vary in two dimensions perpendicular to the background magnetic field. This research confirms that our intuitive understanding of FLR from 1D is a good guide to events in 2D, and places 2D FLR onto a firm mathematical basis by systematic solution of the governing equations. It also reveals the new concept of ‘imprinting’ of spatial forms: spatial variations of the resonant Alfvén wave correlate strongly with the spatial form of the fast wave that drives the resonance.

MI-coupling gives rise to ionosphere-magnetosphere (IM-) waves, and we have made a detailed analysis of these waves for a 1D sheet E-region. IM-waves are characterised by two quantities: a speed  $v_{IM}$  and an angular frequency  $\omega_{IM}$ , for which we have obtained analytic expressions. For an ideal magnetosphere, IM-waves are advective and move in the direction of the electric field with speed  $v_{IM}$ . The advection speed is a non-linear expression that decreases with height-integrated E-region plasma-density, hence, wavepackets steepen on their trailing edge, rapidly accessing small length-scales through wavebreaking. Inclusion of electron inertial effects in the magnetosphere introduces dispersion to IM-waves. In the strongly inertial limit (wavelength  $\lambda \ll \lambda_e$ , where  $\lambda_e$  is the electron inertial length at the base of the magnetosphere), the group velocity of linear waves goes to zero, and the waves oscillate at  $\omega_{IM}$  which is an upper limit on the angular frequency of IM-waves for any wavelength. Estimates of  $v_{IM}$  show that this speed can be a significant fraction (perhaps half) of the  $\mathbf{E}_\perp \times \mathbf{B}_0$  drift in the E-region, producing speeds of up to several hundred metres per second. The upper limit on angular frequency,  $\omega_{IM}$ , is estimated to give periods from a few hundredths of a second to several minutes. IM-waves are damped by recombination and background ionisation, giving an e-folding decay time that can vary from tens of seconds to tens of minutes.

We have also investigated the dynamics and steady-states that occur when the magnetosphere-ionosphere system is driven by large-scale Alfvénic field-aligned currents. Steady-states are dom-

inated by two approximate solutions: an ‘upper’ solution that is valid in places where the E-region is a near perfect conductor, and a ‘lower’ solution that is valid where E-region depletion makes recombination negligible. These analytic solutions are extremely useful tools and the global steady-state can be constructed by matching these solutions across suitable boundary-layers. Furthermore, the upper solution reveals that E-region density cavities form and widen (with associated broadening of the magnetospheric downward current channel) if the downward current density exceeds the maximum current density that can be supplied by background E-region ionisation. We also supply expressions for the minimum E-region plasma-density and shortest length-scale in the steady-state.

IM-waves and steady-states are extremely powerful tools for interpreting MI-dynamics. When an E-region density cavity widens through coupling to an ideal, single-fluid MHD magnetosphere, it does so by forming a discontinuity that steps between the upper and lower steady-states. This discontinuity acts as part of an ideal IM-wave and moves in the direction of the electric field at a speed  $U = \sqrt{v_{IM}^+ v_{IM}^-}$ , which is the geometric mean of  $v_{IM}$  evaluated immediately to the left and right of the discontinuity. This widening speed is typically several hundreds of metres per second. If electron inertial effects are included in the magnetosphere, then the discontinuity is smoothed, and a series of undershoots and overshoots develops behind it. These undershoots and overshoots evolve as inertial IM-waves. Initially they are weakly inertial, with a wavelength of about  $\lambda_e$ , however, strong gradients of  $\omega_{IM}$  cause IM-waves to phase-mix, making their wavelength is inversely proportional to time. Therefore, the waves rapidly become strongly inertial and oscillate at  $\omega_{IM}$ . The inertial IM-waves drive upgoing Alfvén waves in the magnetosphere, which populate a region over the downward current channel, close to its edge. In this manner, the E-region depletion mechanism, that we have detailed, creates small-scale Alfvén waves in large-scale current systems, with properties determined by MI-coupling.

---

## Declaration

---

I, Alexander John Barkway Russell, hereby certify that this thesis, which is approximately 37 000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2007 and as a candidate for the degree of Doctor of Philosophy in September 2008; the higher study for which this is a record was carried out in the University of St Andrews between 2007 and 2010.

Date: 3 Sep 2010.    Signature of Candidate:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 3 Sep 2010.    Signature of Supervisor:

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Access to all of printed copy but embargo on all of electronic publication of thesis for a period of one year on the following ground: publication would preclude future publication.

Date: 3 Sep 2010.    Signature of Candidate:

Signature of Supervisor:

---

## Acknowledgements

---

I would like to thank a number of individuals and organisations for their support, without which I would not have completed this work:

Dr. Andrew Wright has been an ideal supervisor. At all the right times, he has encouraged me, restrained excessive flights of fancy, provided clear insight and allowed me the freedom to explore my own ideas. It has been a pleasure to work together and he has been a very expert guiding hand.

During my time at St Andrews, I have also had the good fortune to work with other members of the Solar and Magnetospheric Theory Group. In particular, I would like to thank Prof. Alan Hood, who provided very useful advice on the steady-state solutions presented in Chapter 4, and introduced me to the asymptotic techniques of boundary-layer matching.

I am grateful for financial support from STFC, in the form of a rolling-grant studentship. I very much hope to see this kind of support continue and grow in the future, so that other individuals will benefit from similar experience. I have also received financial support from ISSI (Bern), whose working group program ‘Multi-Scale Electrodynamics of Magnetosphere-Ionosphere Interactions at High Latitudes’ (led by Prof. Anatoly Streltsov) motivated my research into magnetosphere-ionosphere coupling by field-aligned currents. Further support is acknowledged from KU Leuven and the Solaire Network.

My friends at St Andrews have made these three years a lot of fun. Ruth, Lynsey and Dee are especially important as my office-mates. I also make special mention of those who bake and everyone that’s shared in lunch, frisbee in the park and the dance of eternal joy.

Most importantly, I would like to thank my wife, Fiona, for her ongoing love and support, David, for being a great brother and reminding me of the fun of being outdoors, and my parents, John and Rosalind, who have made this study possible through so many years of love, kindness and encouragement.

---

## Publications

---

The following publications include material from this thesis:

1. A. J. B. Russell and A. N. Wright. Resonant absorption with 2D variation of field line eigenfrequencies. *Astron. & Astrophys.*, 511, A17, 2010. doi:10.1051/0004-6361/200912669.
2. A. J. B. Russell, A. N. Wright, and A. W. Hood. Self-consistent ionospheric plasma density modifications by field-aligned currents: Steady state solutions. *Journal of Geophysical Research (Space Physics)*, 115, A04216, 2010. doi:10.1029/2009JA014836.
3. A. J. B. Russell, A. N. Wright, and A. V. Streltsov. Self-consistent ionospheric plasma density modifications by field aligned currents: time dependent broadening and production of short spatial scales. *In preparation*, 2010.
4. A. J. B. Russell and A. N. Wright. Existence and nature of ionosphere-magnetosphere waves. *In preparation*, 2010.
5. A. J. B. Russell and A. N. Wright. An ionosphere-magnetosphere waves interpretation of intense small-scale Alfvén waves in large-scale field-aligned current systems. *In preparation*, 2010.

---

# Contents

---

<b>Abstract</b>	<b>i</b>
<b>Declaration</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Publications</b>	<b>v</b>
<b>Contents</b>	<b>vi</b>
<b>List of Figures</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Under the Influence of a Star . . . . .	1
1.2 Field-Line Resonance: Coupling the Solar Wind and Magnetosphere . . . . .	2
1.2.1 Strange Events in New and Old England . . . . .	2
1.2.2 Magnetohydrodynamic (MHD) Waves . . . . .	5
1.2.3 What Causes Geomagnetic Pulsations with Minute Periods? . . . . .	6
1.2.4 Field-Line Resonance . . . . .	7
1.2.5 Phase-Mixing . . . . .	9
1.2.6 Field-Line Resonance in Non-Terrestrial Plasmas . . . . .	11
1.2.7 Beyond 1D . . . . .	12
1.3 Magnetosphere-Ionosphere Coupling Via Alfvén Waves . . . . .	13
1.3.1 Our Planet: More Than the Sum of Its Parts . . . . .	13
1.3.2 Connecting Magnetosphere and Ionosphere . . . . .	14
1.3.3 Field-Aligned Currents and Magnetospheric Convection . . . . .	14
1.3.4 Small-Scale Alfvén Waves in Large-Scale Current Systems . . . . .	15
1.3.5 A Series of Valuable Spin-Offs . . . . .	18
<b>2 Field-Line Resonance</b>	<b>20</b>
2.1 Introduction . . . . .	20

2.2	Model . . . . .	20
2.3	Coordinates . . . . .	21
2.4	Computer Simulations . . . . .	22
2.4.1	Governing Equations . . . . .	23
2.4.2	Method . . . . .	25
2.4.3	Results . . . . .	29
2.5	Analytic Solution . . . . .	34
2.5.1	Governing Equations . . . . .	35
2.5.2	Strategy . . . . .	35
2.5.3	Preparation for Series Solution . . . . .	36
2.5.4	Non-Resonant Solution . . . . .	41
2.5.5	Resonant Solution . . . . .	47
2.5.6	Summary of Leading Order Behaviours . . . . .	56
2.6	Discussion . . . . .	58
<b>3</b>	<b>Self-Consistent Treatment of Magnetosphere-Ionosphere Coupling Via Alfvén Waves</b>	<b>60</b>
3.1	Introduction . . . . .	60
3.2	Model . . . . .	60
3.3	Governing Equations . . . . .	62
3.3.1	Coupling of Magnetosphere and Ionosphere . . . . .	62
3.3.2	Magnetospheric Equations . . . . .	63
3.3.3	Single Equation for Ionosphere Coupled to an Ideal Magnetosphere . . . . .	67
3.4	Numerical Code with Ideal Magnetosphere . . . . .	69
3.4.1	1D Governing Equation in Form Solved by Code . . . . .	69
3.4.2	Numerical Methods . . . . .	70
3.4.3	Testing . . . . .	71
3.4.4	Convergence to Steady State . . . . .	72
3.5	Numerical Code with Electron Inertial Effects . . . . .	72
3.5.1	Strategy for Magnetospheric Fields . . . . .	72
3.5.2	Magnetospheric Equations in Form Solved by Code . . . . .	73
3.5.3	Numerical Methods . . . . .	74
3.5.4	Boundary-Conditions Using Ghost-Cells . . . . .	76
3.5.5	Test Case: Normal-Mode Inertial Alfvén Wave . . . . .	78
3.5.6	Incident Inertial Alfvén Wave as an Initial Condition . . . . .	81
3.5.7	Ionospheric Boundary-Condition . . . . .	83

3.6	Discussion . . . . .	85
<b>4</b>	<b>Self-Consistent Steady-States of the Ionosphere-Magnetosphere System</b>	<b>87</b>
4.1	Introduction . . . . .	87
4.2	Computer Simulations . . . . .	87
4.2.1	Setup . . . . .	87
4.2.2	Overview of Steady-States . . . . .	89
4.2.3	Minimum E-Region Plasma-Density . . . . .	90
4.2.4	Shortest Length-Scale in Steady-state . . . . .	91
4.3	Analytic Solutions . . . . .	91
4.3.1	Upper Solution . . . . .	92
4.3.2	Lower Solution . . . . .	94
4.3.3	Critical Points and Integration Constant from Plot of $\Omega b_i(y) - y$ . . . . .	95
4.3.4	Minimum E-Region Plasma-Density . . . . .	95
4.3.5	Plasma-Density Cavity: Existence, Width and Broadening . . . . .	98
4.4	Global Analytic Solution By Boundary-Layer Matching . . . . .	98
4.4.1	Overview of Boundary-Layer Method . . . . .	99
4.4.2	Scaling of Governing Equation . . . . .	101
4.4.3	Scaling and Expanding the Square of the Upper Solution . . . . .	102
4.4.4	Determination of Scaling Parameters $\epsilon$ and $\nu$ . . . . .	103
4.4.5	Boundary-Layer Between Upward and Downward Current-Channels . . . . .	104
4.4.6	Boundary-Layer Farthest From Upward Current-Channel . . . . .	108
4.4.7	Global Solution . . . . .	114
4.4.8	Comparison with Numerical Steady-States . . . . .	115
4.4.9	Shortest Length-Scale in Steady-State . . . . .	115
4.5	Discussion . . . . .	116
<b>5</b>	<b>Existence and Nature of Ionosphere-Magnetosphere Waves</b>	<b>119</b>
5.1	Introduction . . . . .	119
5.2	Ideal IM-Waves . . . . .	119
5.2.1	Advection Equation . . . . .	119
5.2.2	General Properties . . . . .	120
5.2.3	Advection and Damping of a Gaussian Wavepacket . . . . .	122
5.2.4	Physical Cartoon . . . . .	124
5.3	Inertial IM-Waves . . . . .	124

5.3.1	Normal-Mode Analysis . . . . .	124
5.3.2	General Properties and Dispersion Diagram . . . . .	127
5.3.3	Evolution of a Strongly Inertial Wavepacket . . . . .	130
5.4	Multi-Scale Coupling by Wavebreaking of IM-waves . . . . .	132
5.4.1	Wavebreaking of Ideal IM-waves . . . . .	132
5.4.2	Wavebreaking as a Source of Inertial IM-waves . . . . .	135
5.5	Discussion . . . . .	139
5.5.1	Typical Velocities, Frequencies and Decay Times . . . . .	139
5.5.2	IM-Waves and Ionospheric Feedback Instability . . . . .	141
5.5.3	IM-Waves with Fully 2D Sheet E-Region . . . . .	146
5.5.4	IM-Waves and Ionospheric Heaters . . . . .	147
<b>6</b>	<b>Coupled Magnetosphere-Ionosphere Dynamics Driven by Field-Aligned Currents</b>	<b>151</b>
6.1	Introduction . . . . .	151
6.2	Dynamics for ‘Sustainable’ Current Densities . . . . .	151
6.2.1	Numerical Case Study . . . . .	151
6.2.2	Analytic Solution . . . . .	152
6.3	Dynamics for ‘Unsustainable’ Current Densities with Ideal Magnetosphere . . . . .	153
6.3.1	Numerical Case Study . . . . .	153
6.3.2	Analysis In Terms of Ideal IM-Waves . . . . .	155
6.4	Dynamics for ‘Unsustainable’ Current Densities with Electron Inertial Effects . . . . .	157
6.4.1	Overview of Numerical Simulation . . . . .	157
6.4.2	Limitation of $\beta$ in Simulation . . . . .	159
6.4.3	Inertial IM-Waves in Downward Current Channel . . . . .	160
6.5	Discussion . . . . .	162
<b>7</b>	<b>Future Work &amp; Conclusion</b>	<b>166</b>
7.1	Future Work . . . . .	166
7.1.1	Progressing Ionosphere-Magnetosphere Waves . . . . .	166
7.1.2	E-Region Depletion and Small-Scale Waves: Matching Observations . . . . .	167
7.1.3	Amplitude of Small-Scale Alfvén Waves in Large-Scale Current Systems . . . . .	168
7.1.4	Field-Line Resonance . . . . .	168
7.2	Conclusion . . . . .	169
	<b>Bibliography</b>	<b>170</b>

---

## List of Figures

---

1.1	Structure of standing Alfvén waves on Earth’s magnetic field-lines. . . . .	7
1.2	Propagation of fast waves in magnetospheric waveguide. . . . .	8
1.3	Transport of energy from the solar wind to ionosphere by MHD waves. . . . .	8
1.4	The ‘hydromagnetic box’ model. . . . .	9
1.5	A toy model of phase-mixing. . . . .	10
1.6	Field-aligned currents from magnetospheric convection. . . . .	15
1.7	Observations showing intense, short-wavelength electromagnetic disturbances between a large-scale pair of upward and downward field-aligned current channels. . . . .	16
1.8	Simulation of a large-scale current system interacting with the ionosphere. . . . .	17
2.1	Standing Alfvén modes in a hydromagnetic box. . . . .	21
2.2	Orthogonal coordinates defined from contours of field-line eigenfrequency. . . . .	23
2.3	Numerical boundary-conditions mimicking flank of magnetosphere. . . . .	26
2.4	Contour plot of $\tilde{v}_A(\tilde{x}, \tilde{y})$ for sample parameters. . . . .	27
2.5	Cut in $\tilde{v}_A(\tilde{x}, \tilde{y})$ on fixed $\tilde{y}$ . . . . .	27
2.6	Illustration of simulation driver. . . . .	29
2.7	Accumulation of energy-density at a fully 2D resonant contour. . . . .	30
2.8	Phase-mixing of velocity component tangent to the resonant contour. . . . .	31
2.9	Correlation of Alfvénic velocity component and magnetic-pressure gradient. . . . .	32
2.10	Qualitative changes in Alfvén wave energy-density with fast wave driver. . . . .	33
3.1	Diagram of the x-z plane of the MI-coupling model. . . . .	61
3.2	Diagram of the y-z plane of the MI-coupling model. . . . .	62
3.3	Staggered grid used by electron inertial code. . . . .	75
3.4	Use of ghost-cells to implement a symmetric boundary-condition. . . . .	77
3.5	Sample solution from electron inertial code with $\tilde{\tau}_p = 0.2$ (electron plasma oscillations are significant). . . . .	79
3.6	Sample solution from inertial code with $\tilde{\tau}_p = 0.05$ (electron plasma oscillations are negligible). . . . .	79

3.7	Sample solution verifying inertial code matches IAW dispersion relation. . . . .	80
3.8	Saving in simulation runtime obtained by using an initial condition to produce an incident IAW, rather than driving from the top boundary. . . . .	82
4.1	Numerical steady-states for three initial current densities. . . . .	89
4.2	Numerical steady-states for three values of $\beta$ . . . . .	90
4.3	Demonstration that $N_e/N_{min}$ is linear in $\beta$ for the numerical steady-states. . . .	91
4.4	Demonstration that steepest gradient in the numerical steady-state is linear in $\beta$ . .	92
4.5	Sketch of $\Omega b_i(y) - y$ vs. $y$ , from which many key features can be determined. . .	96
4.6	Relationship between upper, lower and global steady-states. . . . .	99
4.7	Comparison of numerical and analytic steady-states. . . . .	115
5.1	Advection of a Gaussian wavepacket under an ideal magnetosphere. . . . .	123
5.2	Cartoon of an ionosphere-magnetosphere wave. . . . .	124
5.3	Dispersion diagram for normal-mode IM-waves. . . . .	129
5.4	Evolution of an IM-wavepacket in the strongly inertial regime. . . . .	131
5.5	Wavebreaking of a non-linear ideal wavepacket. . . . .	133
5.6	Evolution of a large-scale non-linear wavepacket subject to electron inertial effects in the magnetosphere. . . . .	136
5.7	Growth of undershoot behind a non-linear wavepacket subject to electron inertial effects. . . . .	137
5.8	Inertial IM-waves from steepening of non-linear wavepacket: spatial structure. . .	137
5.9	Inertial IM-waves from steepening of non-linear wavepacket: time evolution at fixed location. . . . .	138
5.10	Action of ionospheric feedback instability on an ideal IM-wavepacket. . . . .	143
5.11	IFI simulation by Lysak and Song (2002), also showing IM-waves. . . . .	144
6.1	Evolution of $\tilde{N}$ for a sustainable current density. . . . .	152
6.2	Evolution of $\tilde{N}$ for an unsustainable current density. . . . .	154
6.3	Relationship between discontinuity and the upper and lower steady-states. . . . .	154
6.4	Motion of points $(y, N)$ along characteristic trajectories. . . . .	156
6.5	Evolution of $\tilde{N}$ for an unsustainable current. . . . .	158
6.6	Inertial IM-waves produced by broadening. . . . .	160
6.7	Period of inertial IM-waves produced by broadening. . . . .	161
6.8	Spatial gradients of $\tilde{\omega}_{IM}$ which lead to phase-mixing. . . . .	162

# Introduction

---

## 1.1 Under the Influence of a Star

For centuries, man has believed that life on Earth is influenced by the celestial world. Once, it was considered the domain of the divine, and powerful rulers worked hard to convince their followers that the stars were on their side. Since then, years of careful observation and testing has led to a different world-view, but one in which the Earth is still influenced by events far beyond its surface.

Only one star matters in the modern view: the Sun. It is the warming presence that marks the end of night; it is the source of light and heat upon which nearly all life depends. It is also inconstant, and as human society becomes increasingly interdependent and reliant on technology, we become increasingly vulnerable to the effects of ‘space-weather’, a name given to changes in our environment caused by the changing Sun. Our society began to notice the effects of space-weather in the latter half of the 19<sup>th</sup> Century, when the Industrial Revolution was gathering pace; and it was at this time that space-weather began to take off as a science, as scientists explored links between auroras, geomagnetic storms and sunspot activity. Since then, mankind has expanded both the tools at its disposal and its store of knowledge about the world, but space-weather remains a very active field of research.

Sitting in the outermost layer of the Sun, Earth is constantly assaulted by the solar wind (a low-density plasma rushing away from the Sun at hundreds of kilometres per second) and subjected to large numbers of high-energy photons (including extreme ultraviolet, X-rays and gamma-rays). On Mars, the solar wind has stripped away the atmosphere; on the Moon, radiation would kill all but the hardiest surface life. Fortunately for us, Earth is protected by two defences.

As a first line of defence, Earth’s magnetic field carves out a protective bubble of space called the *magnetosphere*, which electrically charged solar-wind particles sweep around rather than through. As a result, the solar wind passes about 15 Earth radii ( $R_E$ ) above Earth’s surface, a distance that is more than sufficient to protect our atmosphere. This barrier, however, is permeable, and Earth’s magnetic field itself connects the Earth and the solar wind.

Our second protection against the Sun is Earth’s atmosphere: as high energy photons travel towards Earth’s surface, they pass through an increasingly dense atmosphere of neutral particles,

and are absorbed as they ionise these atoms and molecules. Thus, the flux of high-energy photons reaching Earth's surface is greatly reduced — we should still wear sunscreen in bright sun to avoid skin damage, but it is doubtful that multi-cellular life could exist on Earth's surface without the protection of the atmosphere.

Ionisation by solar radiation is greatest in a layer of Earth's atmosphere where the density of neutral particles is high enough to readily absorb incoming photons, and a reasonable flux of high-energy photons remains despite their cumulative removal at higher altitudes. This balance results in Earth's atmosphere having an electrically charged layer called the *ionosphere*, which acts as a transition between our neutral atmosphere and the overlying magnetospheric plasma.

In this thesis, we will consider only the part of the ionosphere that is most significant for electrodynamics. This layer is called the E-region, it sits at an altitude of about 100 km and is approximately 20 km thick. Here, the small collisional cross-section of electrons means that they pass freely through the ions and neutral particles, and any horizontal motion of electrons is uninhibited by collisions. Ions, however, collide regularly with neutral particles, and this causes them (on average) to drift at an angle to the trajectory followed by electrons. Thus, if an electric field causes horizontal movement of electrons and ions in the E-region, then ion-neutral collisions cause a net movement of charge, and the E-region carries a horizontal current. The ability to support a steady electric current with a component perpendicular to the magnetic field is unique to the ionosphere — it occurs in neither the neutral atmosphere (where there are no charged particles to carry a current) nor in the magnetosphere (where steady electric currents are field-aligned).

The central theme running through this thesis is that the Earth, its defences and its space-environment are all part of a single system. In this context, we will focus on two broad topics, examining events where two or more parts of this system interact: (i) the transfer of energy from the solar wind to Earth's magnetic field, in the form of waves; and (ii) interactions between the magnetosphere and ionosphere, which are coupled by electric currents flowing along the magnetic field.

## 1.2 Field-Line Resonance: Coupling the Solar Wind and Magnetosphere

### 1.2.1 Strange Events in New and Old England

On Sunday 20<sup>th</sup> August and Thursday 2<sup>nd</sup> September 1859, the world's skies blazed with auroral light. The ghostly displays lit the Northern Hemisphere as far south as Cuba, and were so bright that the New York Times reported, in Boston “at about one o'clock [in the morning] ordinary print

could be read by the light,” (New York Times, 1859a).

The auroras lit an industrial world: the newspapers of the day, therefore, reported not only the brilliance of the auroral lights, but also the “remarkable effect of the aurora upon the telegraph wires”. A correspondence from telegraph office staff at 31 State Street, Boston, preserves their view of events:

“We have experienced, this morning, a remarkable manifestation of magnetical influence on the wires running in all directions from this office, arising from a magnetic storm, which, were it night, would present a magnificent display of the Aurora. We observed the influence upon the lines at the time of commencing business – 8 o’clock – and it continued so strong up to  $9\frac{1}{2}$  as to prevent any business being done, excepting by throwing off the batteries at each end of the line and working by the atmospheric current entirely! ... The waves were longer than I have ever seen them before, lasting sometimes over a minute; but the same peculiarities of changing of the poles was observed.” (New York Times, 1859b)

This account is an early, but accurate, description of one of the effects of space-weather; namely, that space-weather can cause ultra-low frequency (ULF) magnetic waves with periods of several minutes. In this case the waves produced a periodic electric current in the telegraph wires: when this auroral current worked against an operator’s batteries, the two cancelled out, preventing the sending of any message; when the auroral current and the batteries worked together the result could be dangerous, at least one operator receiving “a very severe electric shock” as “a spark of fire jump[ed] from my forehead”.

The magnetic storms of 1859 were large, but they marked a turning point point in the study of space-weather for two different reasons. The first of these reasons, is that on September 1<sup>st</sup> 1859 Richard Carrington and Richard Hodgson made the first recorded observations of a solar flare, an event that we now know to be intimately connected with space-weather. Their independent accounts appear consecutively in the Monthly Notices of the Royal Astronomical Society (Carrington, 1859; Hodgson, 1859); Carrington gives a longer account, but I find Hodgson’s descriptions more pleasing:

“While observing a group of solar spots on the 1<sup>st</sup> September, I was suddenly surprised at the appearance of a very brilliant star of light, much brighter than the sun’s surface, most dazzling to the unprotected eye, illuminating the upper edges of the adjacent spots and streaks, not unlike in effect the edging of the clouds at sunset; the rays extended in all directions; and the centre might be compared to the dazzling brilliancy of the bright star  $\alpha$  Lyræ when seen in a large telescope with low power. It lasted for some five minutes, and disappeared instantaneously about 11.25 A.M.”

Carrington, after whom the flare came to be named, refrained from suggesting any link between the flare and auroras in print, but it is interesting that the editor chose to add a footnote to his article, which includes the text:

“Mr. Carrington exhibited at the November Meeting of the Society ... and pointed out that ... towards four hours after midnight there commenced a great magnetic storm, which subsequent accounts established to have been as considerable... While the contemporary occurrence may deserve noting, he would not have it supposed that he even leans towards hastily connecting them.”

Despite Carrington’s caution, it is now firmly established that geomagnetic storms are caused by events on the Sun, and, because of this, 1859 is now viewed by many as the year that gave birth to the science of space-weather.

There is one other, important reason for which the magnetic storms of 1859 are remembered, and it is of particular relevance to this thesis: these events were the first time that a sizeable magnetic storm was not only observed, but automatically recorded and graphed in its entirety. These observations were made using “a set of self-recording magnetographs” at Kew Observatory in London, and they were published in a substantial article (Stewart, 1861) that is also notable because (unlike Carrington (1859)) it takes great strides to emphasise the likely link between the auroras and the solar flare. (It also proposes that the upper layers of Earth’s atmosphere are electrically conducting, predicting the ionosphere.) Here, however, we shall focus on what Stewart had to say about magnetic waves:

“The interval of time between two of these minute pulsations may be said to have varied from half a minute, or the smallest observable portion of time, up to four or five minutes.”

“This pulsatory character of the disturbing force agrees well with the nature of its action on telegraph wires, in which observers have notices that the polarity of the current changes very quickly.”

“We have thus, as it were, two sets of waves, the first or smaller of which is superimposed upon the second or larger, just as in the ocean we sometimes see ripples caused by the wind traversing the surface of a great wave.”

We know, now, that the long-period disturbance compared to a great wave (period of hours), was caused by compression of Earth’s magnetosphere as the storm front passed, and by the formation of an energetic ring current around Earth as energetic particles found their way into the magnetosphere from the solar wind. In this thesis, however, it is the smaller oscillations that interest us — the ULF waves with periods of several minutes that Balfour Stewart saw recorded by the

automatic magnetographs, and that telegraph operators experienced first hand in New England. What type of wave could explain these oscillations? And where was the energy coming from? It would be more than a century before both these questions were answered.

### 1.2.2 Magnetohydrodynamic (MHD) Waves

Thanks to the work of Alfvén (1942), we now know that a magnetised plasma, such as the one surrounding Earth, supports *magnetohydrodynamic* (MHD) waves. Here, the term ‘magnetohydrodynamic’ means that the plasma is treated as a conducting fluid, assuming that behaviour of individual protons and electrons is unimportant when considering the bulk behaviour of the plasma (at sufficiently large length-scales and low frequencies). Three independent (decoupled) wave modes exist in a uniform plasma: the *Alfvén* wave, the *fast* magnetoacoustic wave and the *slow* magnetoacoustic wave. In a non-uniform plasma, MHD waves couple together and their behaviour becomes more complex, but the labels Alfvén, fast and slow, and knowledge of their properties, remain valuable tools.

Restoring forces for MHD waves include thermal pressure and the fluid Lorentz force. To visualise the action of Lorentz force in MHD waves, it is useful to divide this force into two components. Using the low-frequency limit of Ampère’s law (appropriate to quasi-neutral plasmas),

$$\mathbf{j} \times \mathbf{B} = \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left( \frac{B^2}{2\mu_0} \right). \quad (1.1)$$

The first term on the right-hand side is the *magnetic tension* force, which acts like tension on a string, with the magnetic field-line taking the place of the string. The second term on the right-hand side acts as a pressure force, so  $B^2/2\mu_0$  is called *magnetic pressure*: if magnetic field is stronger in one region than another, then the magnetic field pushes into the weaker region, trying to equalise magnetic pressure. It is the magnetic tension and pressure forces, together with thermal pressure, that accelerates plasma to create MHD waves.

The first MHD wave that we will consider, and the most widely used in this thesis, is the Alfvén wave. This wave moves plasma perpendicular to the background magnetic field and does not compress the plasma. Here, magnetic tension is the only restoring force. An exact analogue to the Alfvén wave is a wave on a string: on the string, the mass of the string provides inertia, and the restoring force is tension along the string; in a plasma, the mass of the plasma provides inertia, and the restoring force is magnetic tension. Alfvén waves transport energy strictly along the direction of the magnetic field, which means that energy in the form of an Alfvén wave is stuck to a given magnetic field-line; they are non-dispersive; and it can also be shown (e.g. Roberts (1985)) that

the wave speed, called the Alfvén speed, is

$$v_A = \frac{B}{\sqrt{\mu_0 \rho}}, \quad (1.2)$$

where  $\rho$  is the plasma-density.

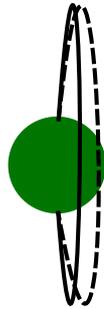
Plasmas also supports two compressive MHD modes called magnetoacoustic waves. These are formally derived in a number of standard textbooks, including a chapter by Roberts (1985) who provides a detailed discussion of their properties; here we shall simply quote the most important results. In the fast wave, the thermal and magnetic pressures are perturbed in phase, creating a wave that travels faster than both the Alfvén speed and the sound speed. The fast wave transports energy in all directions, but attains its greatest speed perpendicular to the magnetic field. In the slow wave, the thermal and magnetic pressure are perturbed out of phase, creating a wave that travels slower than both the Alfvén speed and the sound speed. The slow wave is highly anisotropic and cannot propagate across magnetic field-lines. For magnetospheric work, thermal pressure is often negligible compared to magnetic pressure: in this limit sound speed goes to zero, making the slow wave disappear; meanwhile the fast wave becomes isotropic, travelling in any direction at the Alfvén speed. In this thesis, therefore, we will deal with the Alfvén wave, which transports energy purely along the magnetic field, and the fast wave, which can transport energy in any direction.

### 1.2.3 What Causes Geomagnetic Pulsations with Minute Periods?

It was Dungey (1955) who first connected the work of Alfvén (1942) and the ULF waves observed by Stewart (1861). He proposed that continuous ULF oscillations could be explained as a signature of MHD waves in Earth's magnetosphere. Figure 1.1 illustrates the structure of Alfvén waves standing between conjugate ionospheres (in the Northern and Southern Hemisphere) which act as conducting boundaries. Examining the figure, we can make a number of predictions about pulsations corresponding to standing Alfvén waves, for example: ULF waves should be observed simultaneously at either end of an excited field-line; frequency will generally decrease with increasing latitude (as magnetic field-lines get longer and equatorial field strength decreases); and the waves will be observed in space as well as on the ground. All of these characteristics have since been observed (e.g. Nagata et al. (1963); Sugiura and Wilson (1964); Patel (1965); Samson et al. (1971); Singer et al. (1982)).

Today, it is widely accepted that standing Alfvén waves are the source of continuous geomagnetic pulsations with periods from about 10 seconds to 10 minutes (formally classified as Pc3–5), and they have been observed not only in the terrestrial magnetosphere, but also at five other planets in the solar system (Fraser, 2006). Despite the success of this model, it does raise another

Fundamental Mode



First Overtone



Figure 1.1: Dungey (1955) proposed that a wide range of geomagnetic pulsations could be explained as standing Alfvén waves on Earth’s magnetic field-lines. Here, field-line displacements corresponding to the fundamental mode and first overtone are illustrated.

question: if standing Alfvén waves trap energy on magnetic field-lines, how are they excited in the first place?

#### 1.2.4 Field-Line Resonance

The Alfvén waves that cause geomagnetic pulsations must have an energy source, and an important clue to their origin is an observed symmetry between Earth’s dawn and dusk flanks: on both flanks the phase-speed of Pc3–5 ULF waves is typically away from the Sun, and the waves disappear at noon (Anderson et al., 1990; Samson et al., 1992). This suggests that the solar wind, which travels away from the Sun, is somehow providing the required energy.

For energy to be transferred from the solar wind to standing Alfvén waves inside the magnetosphere, that energy must be transported across Earth’s magnetic field. Recalling the properties of MHD waves (Section 1.2.2), fast magnetoacoustic waves are capable of providing this transport, and are the only type of MHD wave to do so in a cold plasma. Figures 1.2 and 1.3, therefore, show ingredients of a plausible model: the solar wind excites fast waves in the outer magnetosphere, which do not penetrate as far as the ionosphere because they refract and become evanescent in the inner magnetosphere; ionospheric disturbances can be caused, however, if it is possible for fast waves to transfer part of their energy to resonant Alfvén waves, which transport energy along the background magnetic field.

The process connecting fast and Alfvén waves is called *field-line resonance*, and is the main focus of Chapter 2. In his work identifying ULF pulsations with Alfvén waves, Dungey (1955) noted that fast and Alfvén waves decouple if the waves are invariant along an invariant direction of the equilibrium, e.g. in a dipole magnetic field with an axisymmetric density profile, MHD waves decouple if their perturbations are axisymmetric. If such a symmetry is absent, then (in a cold

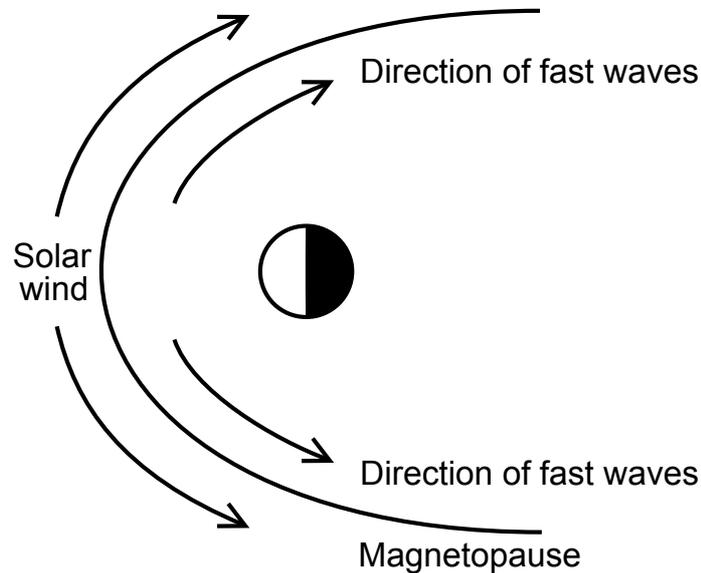


Figure 1.2: The solar wind excites fast waves in the outer magnetosphere, which propagate tailwards along the magnetospheric waveguide.

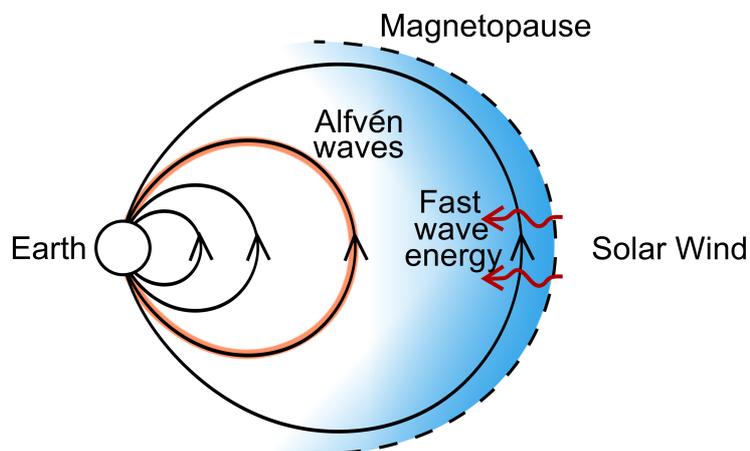


Figure 1.3: Fast waves, excited by the solar wind, transport energy across the background magnetic field, allowing this energy to penetrate the outer magnetosphere. Due to increases in Alfvén speed with proximity to Earth, these waves are refracted away from the inner magnetosphere, becoming evanescent there. Thus it first appears that energy in fast waves is incapable of reaching Earth. It is possible, however, for field-line resonance to transfer part of the energy in the fast waves to resonant Alfvén waves, which transport energy along the background magnetic field and cause ionospheric disturbances.

plasma for which the slow mode disappears) we are presented with two coupled wave equations, one for fast waves and one for Alfvén waves.

The most influential works on field-line resonance are probably Southwood (1974) and Chen and Hasegawa (1974) (see also Radoski (1971)). In order to investigate coupling in the simplest useful configuration, Southwood used a model known as the ‘hydromagnetic box’, which replaces Earth’s curved magnetic field with a straight, uniform magnetic field, and varies Alfvén speed through density (Figure 1.4). Southwood varied density in the  $x$ -direction only, producing a 1D variation of field-line Alfvén eigenfrequencies, and obtained an analytic solution by assuming normal-modes in  $y, z$  and time.

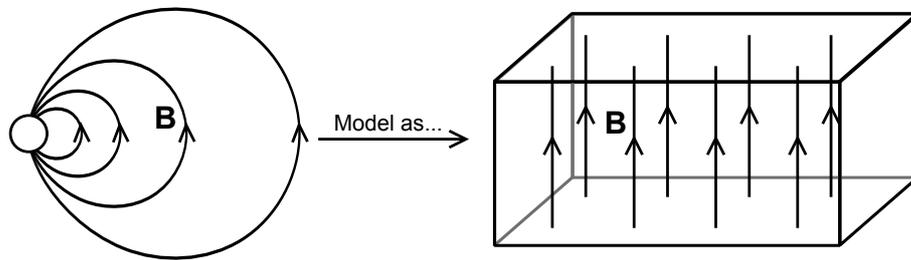


Figure 1.4: The ‘hydromagnetic box’ model, in which Earth’s magnetic field is replaced with a straight, uniform magnetic field. Top and bottom boundaries are chosen to mimic the ionosphere and gradients in field-line eigenfrequency are obtained by varying density perpendicular to the magnetic field.

The coupled solution thus obtained has a singular Alfvén wave component where the field-line Alfvén frequency  $\omega_A(x)$  matches the normal-mode frequency  $\omega$ . This is interpreted as a flow of energy from the fast wave to the resonant Alfvén wave, with a corresponding jump in fast-wave Poynting flux at the resonance. Thus, field-line resonance takes energy from low-amplitude fast waves, and gradually builds up a large-amplitude, resonant Alfvén wave (a principal result that has since been confirmed by time-dependent computer simulations, e.g. Allan et al. (1986); Lee and Lysak (1989); Rickard and Wright (1994); Wright and Rickard (1995); Mann et al. (1995) and other references therein).

### 1.2.5 Phase-Mixing

Whenever a solution oscillates with a spatially-dependent frequency, nearby parts get increasingly out of phase, leading to the appearance of a diminishing length-scale that separates points exactly one cycle out of phase. This process is called *phase-mixing*, and produces a characteristic length-scale is known as the *phase-mixing length*,  $L_{ph}(x, t)$ .

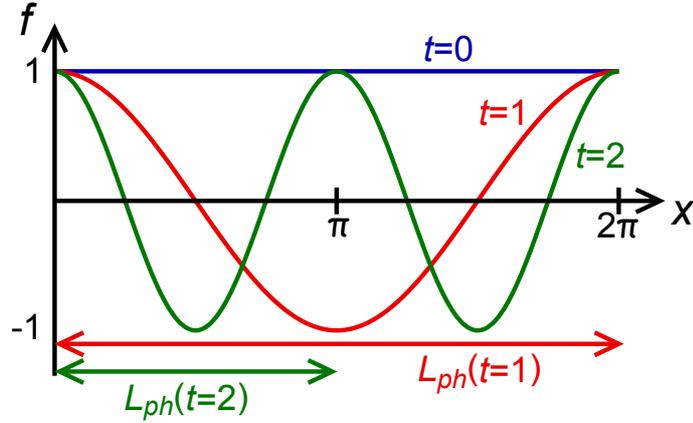


Figure 1.5: A toy model of phase-mixing. Due to spatial variation of frequency, points on an oscillating function drift out of phase. As phases mix, a spatial wavelength develops, called the phase-mixing length,  $L_{ph}$ . Note that  $L_{ph}$  gets shorter in time.

As an example, Figure 1.5 shows how the function

$$f(x, t) = \cos(\omega(x)t) \quad (1.3)$$

evolves in time for

$$\omega(x) = 2\pi + x. \quad (1.4)$$

At  $t = 0$ ,  $f = 1$  for all  $x$ ; the function is uniform so the phase-mixing length is infinite. When  $t = 1$ , the function has completed one full cycle of oscillation at  $x = 0$ , and two cycles have been completed at  $x = 2\pi$ ; thus,  $L_{ph}(t = 1) = 2\pi$ . At  $t = 2$  (the final time represented in Figure 1.5), the function has completed two full cycles of oscillation at  $x = 0$ , and three full cycles have been completed at  $x = \pi$ , giving  $L_{ph}(t = 2) = \pi$ . Phase-mixing continues to have this effect, creating smaller and smaller length-scales as time increases.

It is reasonably straightforward to obtain a formula that estimates  $L_{ph}(x, t)$  for general  $\omega(x)$  (e.g. Mann et al. (1995)). To begin, let us define  $n(x, t)$  as the number of cycles of oscillation completed at  $(x, t)$ :

$$n(x, t) = \frac{t}{T(x)} = \frac{\omega(x)t}{2\pi}. \quad (1.5)$$

The phase-mixing length can then be estimated from the difference in  $n$  between two points separated by a distance  $\Delta x$ , thus

$$|n(x + \Delta x, t) - n(x, t)| = \frac{\Delta x}{L_{ph}(x, t)}. \quad (1.6)$$

Substituting for  $n$  using 1.5,

$$|\omega(x + \Delta x) - \omega(x)| = \frac{2\pi\Delta x}{L_{ph}(x, t)t}. \quad (1.7)$$

If  $L_{ph}$  is independent of position, then putting  $\Delta x = L_{ph}(t)$  into equation (1.6) shows that two points separated by a distance  $L_{ph}$  are exactly one cycle out of phase, as expected from the definition of  $L_{ph}$ . More generally, however, we can solve equation (1.7) in the limit  $\Delta x \rightarrow 0$  to obtain  $L_{ph}(x, t)$  at a point; doing so, the left-hand side can be Taylor expanded to first order, and some rearranging gives

$$L_{ph}(x, t) = \frac{2\pi}{|d\omega/dx|t}. \quad (1.8)$$

Therefore, the phase-mixing length is inversely proportional to time.

It is natural that phase-mixing should be discussed alongside field-line resonance, because field-line resonance excites Alfvén waves on a gradient of field-line Alfvén eigenfrequency. Neighbouring Alfvén waves, therefore, will phase-mix, producing small length-scales transverse to the background magnetic field. It is worth bearing in mind that phase-mixing will occur at any frequency gradient, and we shall encounter one other example of phase-mixing in this thesis.

### 1.2.6 Field-Line Resonance in Non-Terrestrial Plasmas

Field-line resonance is a fundamental plasma process that requires only a gradient of field-line eigenfrequency and a source of compressive (fast wave) energy; as such, it is applicable in many different situations. At present, the concept is very much in vogue in solar physics, where it is called *resonant absorption*. Interest in the solar community was provoked by observations of coronal loop ‘kink’ oscillations: the swaying of dense loops, excited, for example, when the loop is struck by a blast wave from a solar flare. Kink oscillations were predicted by Edwin and Roberts (1983), and first observed by Aschwanden et al. (1999) and Nakariakov et al. (1999). The first observations sparked a flurry of theoretical research, because the oscillations decayed with unexpected rapidness (decay time of about 15 mins, compared with a period of about 5 mins).

Several decay mechanisms have been proposed, and field-line resonance is a leading candidate (e.g., Hollweg & Yang 1988; Goossens et al. 1992; Ruderman & Roberts 2002; Terradas et al. 2006 and references therein). In a coronal loop, kink waves cause a weak compression of the magnetic field, which allows them to act like a fast wave and resonantly couple to a Alfvén waves. Furthermore, the frequency of a kink wave lies between the Alfvén frequency at the heart of a loop, and the Alfvén frequency of the loop’s environment; hence, for a continuous profile of Alfvén frequency, there is always a resonant surface at which waves couple. It is therefore

unavoidable that field-line resonance removes energy from the kink wave (which is observed because it displaces the axis of its loop), and transfers it instead to Alfvén waves in the boundary of the loop (which do not displace the loop axis). Therefore, the observed oscillation decays. Of the various decay mechanisms that have been proposed to date, field-line resonance is notable for its inevitability, and for being the only mechanism to reproduce the observed damping time without invoking anomalous values for physical parameters.

Field-line resonance continues to be applied to increasingly distant objects, one of the more recent examples being the work of Reznia and Samson (2005), who turned their attention to the origin of periodic oscillations in X-ray flux from an accreting neutron star. They suggested that the X-ray flux is modulated by Alfvén waves in the neutron star's magnetosphere, which in turn draw their energy, through field-line resonance, from fast waves caused by non-steady-accretion flows. Field-line resonance is therefore an attractive topic to study, because it remains at the forefront of a wide and growing range of astrophysical research.

### 1.2.7 Beyond 1D

The work of Southwood (1974) and Chen and Hasegawa (1974) has been extremely valuable, establishing the basis of field-line resonance, but their work is based on a 1D model. So, one might ask, how confident are we that field-line resonance is not just an artefact of this 1D model? Or, if the core idea is robust, what new behaviours emerge in a more general model? In relaxing the geometric constraints of the original model, there are two complementary issues to address: (i) we should like to know how field-line resonance works for generalised field-line Alfvén eigenmodes, as will occur if Alfvén speed varies along the background magnetic field as well as across it; and (ii) we should also like to know how field-line resonance works if field-line Alfvén eigenfrequencies vary in two dimensions, as will occur if Alfvén speed varies in the two dimensions perpendicular to the background magnetic field.

Analytic efforts have already had some success. Notably, Schulze-Berge et al. (1992) allowed the density profile to vary in three dimensions: they obtained a leading-order solution, and showed that this is singular at the resonant surface, but stopped short of the full solution. In contrast, Thompson and Wright (1993) systematically obtained a full solution for a less general model, where density varied in one dimension across the background magnetic field and a second dimension along it. Their complete solution is singular where field-line Alfvén eigenfrequency matches the driving frequency, whatever the Alfvén eigenmode may be. They also demonstrated the existence of an overlap integral that determines the efficiency of the excitation: excitation is most efficient when the spatial form of the driver matches the eigenmode of the resonant field-lines. We can, therefore, think of their study as addressing the first issue identified above.

The structure of the magnetosphere and of coronal loops encourages us to investigate equilibria that vary in two dimensions across the background magnetic field (so that field-line Alfvén eigenfrequencies vary in two dimensions). In the magnetosphere, Alfvén speed is not axisymmetric for the dawn flank, magnetotail or dusk flank, although observations show that field-line resonance persists in these regions (Anderson et al., 1990; Takahashi and Anderson, 1992; Takahashi et al., 1996). Similarly, coronal loops may have non-circular cross-section (Ruderman, 2003), and there is also good evidence that loops have substructure, consisting of elemental magnetic flux strands of widths less than 2 Mm (Schmelz et al., 2001, 2003, 2005; Schmelz, 2002; Martens et al., 2002; Aschwanden and Nightingale, 2005; Aschwanden, 2005).

Two numerical simulations (Lee et al., 2000; Terradas et al., 2008) have already confirmed that field-line resonance persists when field-line Alfvén eigenfrequencies vary in 2D: both of these simulations considered a density profile that varied in two dimensions across the background magnetic field, and sent in a pulse of fast wave energy; this energy was deposited at locations where field-line Alfvén eigenfrequency matched a global fast wave eigenfrequency of the whole system, consistent with field-line resonance.

Despite the achievements of these previous works, there is little analytic work modelling field-line resonance with 2D field-line Alfvén eigenfrequencies. We have, therefore, made considerable efforts to explore this problem, the results of which are presented in Chapter 2. Here, continuously driven numerical simulations confirm that field-line resonance is robust and demonstrate new features arising from the 2D eigenfrequencies. We also present a full analytic solution, which achieves the important task of setting field-line resonance onto a firm, mathematical foundation when eigenfrequencies vary in 2D.

## 1.3 Magnetosphere-Ionosphere Coupling Via Alfvén Waves

### 1.3.1 Our Planet: More Than the Sum of Its Parts

Very often, we find that systems are more than the sum of their parts, exhibiting behaviour that cannot be understood by considering their parts in isolation. Examples of such ‘emergent behaviour’ are found in diverse subject areas, examples being the strength of alloyed metals, the harmony of a musical chord, and collective behaviour of ants. Therefore, a successful scientific approach tries to understand our world not only by breaking it into its simplest parts, but by subsequently rebuilding the whole, pausing to understand every hierarchy of interaction. Earth is no different, and if we wish to truly understand our planet, then we need to understand how its different parts act together, including the magnetosphere and ionosphere. This need to understand the magnetosphere and ionosphere as a single system forms the second focus of this thesis: addressing this

need has revealed interesting surprises, as unexpected behaviour emerges from the combination of simpler parts.

### 1.3.2 Connecting Magnetosphere and Ionosphere

A sizeable portion of ionospheric and magnetospheric research, both past and ongoing, treats one part of the system passively, i.e. by assuming  $A$  influences  $B$ , but  $B$  does not influence  $A$ . This simplification is valid if one assumes that part of the system does not change on the time-scale of interest, or that the variation can be prescribed by some predetermined function. The field-line resonance study presented in Chapter 2, for example, uses a passive ionosphere, examining events in the magnetosphere under the assumption that the ionosphere is highly conducting during the time of interest. Other problems have been studied assuming a passive magnetosphere, for example, research into the auroral ionosphere has sometimes treated the ionosphere as the only active component, choosing to specify a boundary-condition with the magnetosphere (e.g. Karlsson and Marklund (1998)).

An approach that treats half of the system passively has limits, as the following example from Cran-McGreehin et al. (2007) shows. If we assume that the ionosphere is always highly conducting, then we know that a standing magnetospheric Alfvén wave has anti-nodes in magnetic field perturbation and current density at its ionospheric footpoints. Under these conditions, the downward current that such a wave draws from the ionosphere can easily reach  $1\mu\text{Am}^{-2}$ , and will be sustained for a half wave-period, lasting up to several minutes. This corresponds to an outflow of  $6.2 \times 10^{12}$  electrons  $\text{m}^{-2}\text{s}^{-1}$  from the ionosphere, some ten times greater than the rate at which electrons are produced there during night. Neglecting production of electrons in comparison to the outflow, the downward current will consume a night-side ionosphere (with height-integrated number density of  $2 \times 10^{14}\text{m}^{-2}$ ) in about 30 seconds, well within the lifetime of the downward current. This, of course, invalidates the original assumption that the ionosphere is perfectly conducting at all times. To solve the problem properly (self-consistently) we must include not only the effects of the magnetosphere on the ionosphere, but also the corresponding effects of the changing ionosphere on the magnetosphere.

### 1.3.3 Field-Aligned Currents and Magnetospheric Convection

The persistent presence of polar auroras, which are excited by energetic electrons moving along Earth's magnetic field, attests the existence of a large-scale current system flowing through the magnetosphere and ionosphere. These currents occur even when the system is unchanging in time (hence ULF waves are absent) and are primarily driven by convection of magnetospheric plasma, with energy ultimately coming from the solar wind.

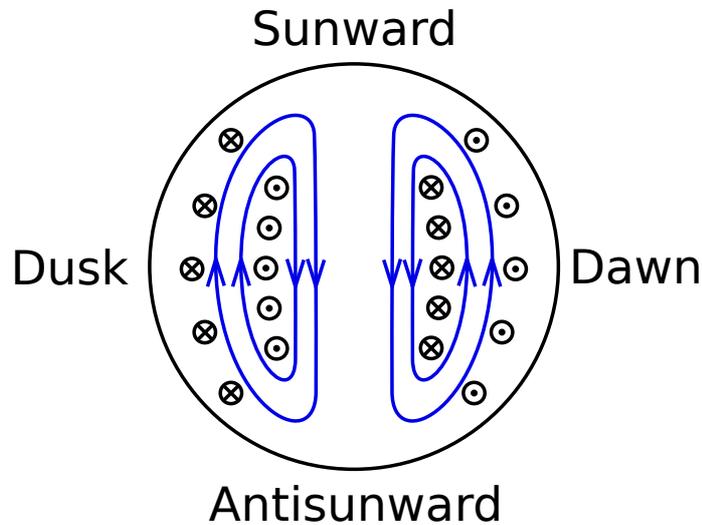


Figure 1.6: The antisunward flow of the solar wind causes magnetospheric convection, which is away from the Sun at high latitudes (where magnetic field-lines connect to the solar wind), and towards the Sun at lower latitudes (where magnetic field-lines are closed). Here, streamlines of the flow are sketched blue. Ionospheric drag, which resists the magnetospheric flow, causes a horizontal magnetic field, which is oppositely directed to the flow in the Northern Hemisphere. Thus, the flow vortices are also magnetic vortices, giving rise to field-aligned currents (shown here by the black circles).

Magnetospheric convection arises from the fact that magnetic field-lines over Earth's magnetic polar caps typically join to magnetic field-lines in the solar wind. As a result, they are pulled over the poles in an antisunward direction, as the interplanetary part of the field-line moves with the solar wind. This part of the process transports magnetic flux from the nose of the magnetosphere to the tail of the magnetosphere. In order that the magnetic flux at the nose be replenished, there must be a return flow in the sunward direction, and this happens at lower latitudes, where field-lines are closed (i.e. do not join with the solar wind). In each hemisphere, the resulting flow pattern forms twin vortices, and Figure 1.6 illustrates these for the Northern Hemisphere. The ionosphere causes a drag at the base of magnetospheric field-lines, so the convection flow produces a horizontal magnetic field perturbation, which, is opposite to the flow for the Northern Hemisphere. The flow vortices, therefore, correspond to magnetic vortices, and are threaded by large-scale field-aligned currents, referred to as Region 1 currents. There is a second current system at low-latitudes, caused by the low-latitude magnetic shear, and these are called Region 2 currents. The whole of this convection-driven system typically carries a few mega-amperes of current.

### 1.3.4 Small-Scale Alfvén Waves in Large-Scale Current Systems

In deciding to study the magnetosphere and ionosphere as one system, we are motivated by more than curiosity alone (although this should always be reason enough). More important than any

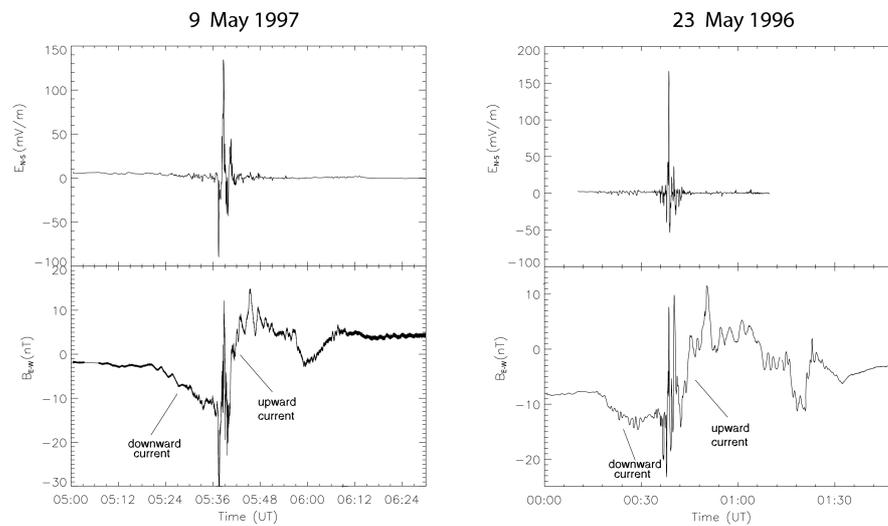


Figure 1.7: North-South electric field and East-West magnetic field measured by the Polar satellite on 9 May 1997 (left) and 23 May 1996 (right). Both sets of observations show intense, short-wavelength electromagnetic disturbances between a large-scale pair of upward and downward field-aligned current channels. Adapted from Keiling et al. (2005). Copyright 2005 American Geophysical Union. Modified by permission of American Geophysical Union.

other reason, we study it because nature challenges us to do so.

In recent decades, satellites orbiting Earth have observed a new class of Alfvén wave, whose origin has defied explanation in terms of either the ionosphere or magnetosphere alone. Figure 1.7 (extracted from Keiling et al. (2005) with some modifications) shows observations from the Polar spacecraft that illustrate the typical conditions under which these waves are seen. In the two events shown, Polar flew through a large-scale pair of upward and downward field-aligned current channels, and it encountered an intense, short-wavelength electromagnetic disturbance between the two current channels. Similar observations have been made using both the FAST (Paschmann et al., 2002) and Cluster satellites (e.g. Karlsson et al. (2004); Johansson et al. (2004)), revealing that the small-scale disturbances normally lie just inside the downward current channel. Since Cluster is a constellation of four separate spacecraft, it has been possible to separate spatial and temporal variation in the Cluster data, and the typical period of these waves has been estimated as about 20–40 seconds.

The period of these waves is too long for a traditional ionospheric explanation, such as trapping inside an ionospheric Alfvén resonator. In addition to this, they are observed at an altitude of 4–7  $R_E$  which is well above any ionospheric trapping region. It is therefore very unlikely that these waves are of an exclusively ionospheric origin. On the other hand, the period is much shorter than typical magnetospheric time-scales, such as the Alfvén wave transit time from one hemisphere to another. Furthermore, these waves frequently occur near the boundary between open and closed

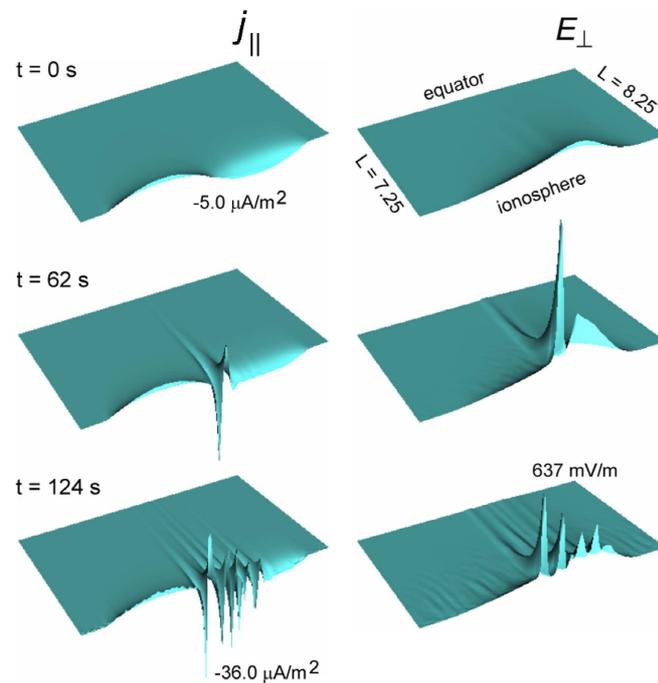


Figure 1.8: Interaction between a large-scale current system and the ionosphere, modelled by Streltsov and Lotko (2004). The interaction produces intense, small-scale waves between the large-scale upward and downward current channels, qualitatively matching satellite observations. Copyright 2004 American Geophysical Union. Modified by permission of American Geophysical Union.

field-lines, which makes it hard to identify them with any plausible magnetospheric cavity. Thus, attempts to match the period of these waves with either the magnetosphere or ionosphere (in isolation) have proved fruitless.

The one promising line of research has been the hypothesis that small-scale waves are generated by a non-linear interaction of the large-scale, magnetospheric field-aligned currents with an active ionosphere. The plausibility of such an explanation has been firmly established via computer simulations performed by Streltsov and Lotko (2004, 2005); Streltsov and Karlsson (2008). Here, the authors used an ‘all-inclusive’ computer code that included two-fluid effects, dipole geometry and ionospheric substructure (E-region and F-region). Their results (see Figure 1.8) show that intense small-scale electromagnetic waves are indeed produced at the boundary between upward and downward current channels, by the interaction between magnetosphere and ionosphere, and the simulated waves bear a strong resemblance to observations. Whilst these simulations represent significant progress by showing ionosphere-magnetosphere interaction is crucial to the formation of these waves, they have not produced the desired detailed explanation, due in part to the large amount of physics contained in the model.

Research into these waves received added impetus in 2008 when the International Space Sci-

ence Institute (ISSI), in Bern, awarded support to a working group to progress understanding of these waves. Without this support, and Anatoly Streltsov (who led the team) drawing our attention to this work, it is unlikely that we would have joined efforts to solve this problem, and we are grateful to them both for starting this involvement. The original proposal (Streltsov, 2008) defined two central questions that the working group would try to answer:

- 1) What mechanism generates small-scale waves observed in the magnetosphere?
- 2) What defines the frequency, amplitude and transverse scale-sizes of these waves?

I hope to offer some answers to these questions within this thesis.

### 1.3.5 A Series of Valuable Spin-Offs

In trying to explain the origin of small-scale Alfvén waves in large-scale current systems, we have been forced to ask some fundamental questions about interactions between the ionosphere and the magnetosphere. In Chapter 3 we set out models of the coupled system, which we have kept as simple as possible for three reasons: (i) a simple model allows us to focus on the most fundamental behaviour; (ii) a simple model is more tractable than one that includes (unessential) additional physics; and (iii) simple models are often the most robust, since their behaviour relies on the most fundamental of principles. Having established suitable models, we have then addressed a number of questions which are prerequisites to our main goal, but are also of significant importance in their own right.

Chapter 4 examines steady-states of the coupled magnetosphere-ionosphere (MI) system. As well as preparing for a discussion of dynamics, this also produces results of great relevance to plasma-density cavities in the E-region layer of the ionosphere. Previous studies have established that E-region plasma-density cavities form in regions of strong, downward field-aligned current and that the downward current channel widens as the cavity broadens to its final width (Doe et al., 1995; Blixt and Brekke, 1996; Karlsson and Marklund, 1999; Marklund et al., 2001; Aikio et al., 2002, 2004; Cran-McGreehin et al., 2007; Michell et al., 2008). We have successfully found a general method by which the final steady-state can be constructed, significantly extending previous work, and allowing us to find the final width and minimum density of a density cavity, as well as the finest length-scale in the steady-state. Furthermore, we have also quantified the conditions under which an E-region density cavity forms and the downward-current channel broadens; generalising a major result from Cran-McGreehin et al. (2007).

Chapter 5 progresses to a first look at dynamics of the coupled MI system. Here, we consider a type of wave driven by interplay between height-integrated E-region number density and field-aligned currents in the magnetosphere. Although the linear limit of these waves has been

recognised as central to the ionospheric feedback instability, we give a fresh perspective, using a model that does not include reflection of upgoing Alfvén waves back towards the E-region. This simplification permits a much fuller investigation than previous studies, including consideration of non-linear effects, and provides substantial appreciation of the underlying physics. Since the waves originate from coupling between the ionosphere and magnetosphere, I have (for now) simply dubbed them ionosphere-magnetosphere waves, or IM-waves for short. When IM-waves have large horizontal spatial-scales (electron inertial terms negligible in the magnetosphere), they are advective and disturbances move in the direction of background electric field. For large-amplitude waves, wave speed is a function of E-region plasma-density, making these waves non-linear and subject to wavebreaking. When small spatial-scales make electron inertial terms significant in the magnetosphere, the nature of the waves changes: as length-scales make the transition into the inertial regime, the waves become dispersive; once they are strongly inertial ( $k_{\perp} \lambda_e > 2\pi$ ), the waves' group velocity goes to zero and they remain in a fixed location, oscillating at a characteristic frequency. IM-waves are at once an interesting phenomenon, and an extremely useful tool with which to understand the coupled dynamics of magnetosphere and ionosphere. They also suggest intriguing applications, such as the possibility to efficiently send waves into the magnetosphere by appropriate driving of the ionosphere.

Having made this preparation, we are able to properly investigate the evolution of a coupled ionosphere and magnetosphere, driven by large-scale field-aligned current in the form of an incident Alfvén wave. Results are presented in Chapter 6. First, we investigate the dynamics of a weak current system, seeing that for low current densities the ionosphere can play a more-or-less passive role. Having understood this case, we then examine dynamics for a current density strong enough to form an E-region plasma-density cavity. Here, we see how a plasma-density cavity forms and broadens, using our knowledge of steady-states and IM-waves to explain the process. For the case of an ideal magnetosphere, broadening forms a discontinuity in E-region number density, by a means similar to wavebreaking of ideal IM-waves. This forces us to include electron inertia in the magnetosphere, and when we do, we find that broadening produces inertial IM-waves between the two current channels, radiating small-scale Alfvén waves into the magnetosphere.

# Field-Line Resonance

---

## 2.1 Introduction

Field-line resonance (FLR) is a fundamental MHD process that transfers energy from fast magnetoacoustic waves to Alfvén waves, thus combining their properties to open new routes for energy to flow through a system. This concept was introduced in Chapter 1, where we considered examples of FLR in the terrestrial magnetosphere, the solar corona and in neutron star magnetospheres. We also emphasised the need for new research to investigate the properties of FLR when field-line Alfvén eigenfrequencies vary in 2D.

This chapter addresses the need for 2D understanding. First, time-dependent computer simulations are presented that explore field-line resonance in an appropriate 2D setup. These are an excellent test of our intuitions, allowing us to assess how well they carry over from 1D to 2D. They also provide an opportunity to explore new features that arise in the 2D problem. Second, we set the 2D problem onto a firm mathematical foundation. This mathematical solution is a prerequisite to complete understanding of 2D field-line resonance, and we are able to tease new features from it.

## 2.2 Model

The ‘hydrodynamic box’ model used by Southwood (1974) is easily adapted to study field-line resonance when field-line Alfvén eigenfrequencies vary in 2D. As in the original model (see Section 1.2.4), Earth’s magnetic field is replaced with a simpler magnetic field that is both straight and uniform; gradients in field-line eigenfrequency are obtained by varying density across the background magnetic field; and top and bottom boundaries are chosen to mimic a highly conducting ionosphere by setting  $\mathbf{E}_\perp = 0$  at these surfaces. Figure 1.4 illustrates the model.

A straight, uniform magnetic field ensures that magnetic pressure and tension forces are zero in the equilibrium. Furthermore, equilibrium density gradients are maintained by the assumption that the plasma is cold, so that there is zero plasma pressure (this has the additional advantage of setting the slow magnetoacoustic speed to zero, eliminating the slow wave from later considerations).

Gravity is neglected and we limit ourselves to times at which dissipative and non-linear effects are unimportant.

In our model, the direction of the equilibrium magnetic field defines the  $z$ -axis, i.e.  $\mathbf{B}_0 = B_0 \hat{z}$ . Density,  $\rho_e(x, y)$ , is a function of two dimensions only, so that density varies across the magnetic field, but not along it. This is the simplest equilibrium that provides 2-dimensional field-line eigenfrequencies. Under these constraints, field-line Alfvén eigenfunctions are Fourier modes with an  $\exp(\pm i k_z z)$  dependence (Figure 2.1). In an infinite medium, any value of  $k_z$  can be considered, but given perfectly conducting ionospheric boundary-conditions,  $k_z$  is quantised so that  $k_z = n\pi/L$ , where  $L$  is the length of the field-line and  $n$  is an integer. The corresponding field-line eigenfrequency is  $\omega_A(x, y) = k_z v_A(x, y)$ , where  $v_A(x, y) = B_0 / \sqrt{\mu_0 \rho_e(x, y)}$  is the equilibrium Alfvén speed.

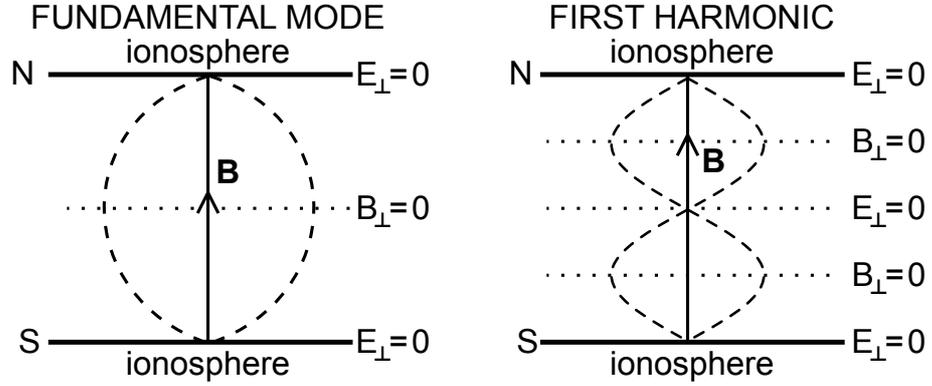


Figure 2.1: The fundamental standing Alfvén mode and first overtone in a hydromagnetic box with highly conducting ionospheres. In each part of this figure, the solid line with an arrow represents the equilibrium position of the field-line and the dashed lines show the field-line's maximum displacement. Displacement antinodes are antinodes for the electric field perturbation and nodes for the magnetic field perturbation. Displacement nodes (as occur at perfectly conducting boundaries) are nodes for the electric field perturbation and antinodes for the magnetic field perturbation.

Matters are further simplified by matching the  $z$ -dependence of our driver to a specific Fourier mode. Since Fourier modes are orthogonal, this means that there is exactly one eigenfunction (in  $z$ ) that can be driven on any given field-line. We therefore follow a normal-mode analysis.

## 2.3 Coordinates

In the work that follows, it is convenient to use two coordinate systems, and to change between these. In each system, the three coordinate directions are orthogonal to one another.

The first coordinate system is Cartesian coordinates  $(x, y, z)$ , in which a coordinate unit vector has fixed orientation everywhere in space. We have already made one requirement on this system,

aligning  $\hat{\mathbf{z}}$  with the background magnetic field.

The second coordinate system uses curvilinear coordinates  $(X, Y, z)$ , illustrated in Figure 2.2. The unit vector  $\hat{\mathbf{z}}$  remains aligned with the background magnetic field as before, but this time  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{Y}}$  are defined from isosurfaces of  $\omega_A(x, y)$ , setting  $\hat{\mathbf{X}}$  perpendicular to isosurfaces of  $\omega_A$  and  $\hat{\mathbf{Y}}$  tangent to isosurfaces of  $\omega_A$ . This is done so that  $\omega_A = \omega_A(X)$  in the new coordinate system, i.e. field-line eigenfrequency is a function of one variable only,  $X(x, y)$ . Scale factors for  $X$  and  $Y$ , which we label  $h_X(X, Y)$  and  $h_Y(X, Y)$  respectively, may be different to unity and are generally 2D scalar functions.

For the equilibrium considered here, the unit vectors of  $(X, Y, z)$  are

$$\hat{\mathbf{z}} = \frac{\mathbf{B}_0}{B_0}, \quad (2.1)$$

$$\hat{\mathbf{X}} = \frac{\nabla \rho_e}{|\nabla \rho_e|}, \quad (2.2)$$

$$\hat{\mathbf{Y}} = \frac{\mathbf{B}_0 \times \nabla \rho_e}{B_0 |\nabla \rho_e|}. \quad (2.3)$$

There is potential for trouble where  $|\nabla \rho_e| = 0$  (i.e. where density is locally uniform) because  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{Y}}$  are indeterminate at such locations; we shall restrict our use of these coordinates to regions where this problem does not occur.

Our final consideration in establishing the coordinates  $(X, Y, z)$  is to choose a useful origin for  $X$ . If the system is driven at a frequency  $\omega_d$  that somewhere matches the natural Alfvén frequency,  $\omega_A(X)$ , then the surface on which these frequencies match is called the *resonant layer*. In situations where this resonance occurs, it is helpful to define the origin of  $X$  so that  $X = 0$  on the resonant surface. This completes the definition of  $(X, Y, z)$ .

## 2.4 Computer Simulations

Computer simulations are often a very good way to begin an investigation: they allow us to simply watch what happens, reveal areas where preconceived ideas fail, and guide us toward new theories. With this in mind, we take our first look at 2D field-line resonance through simulations.

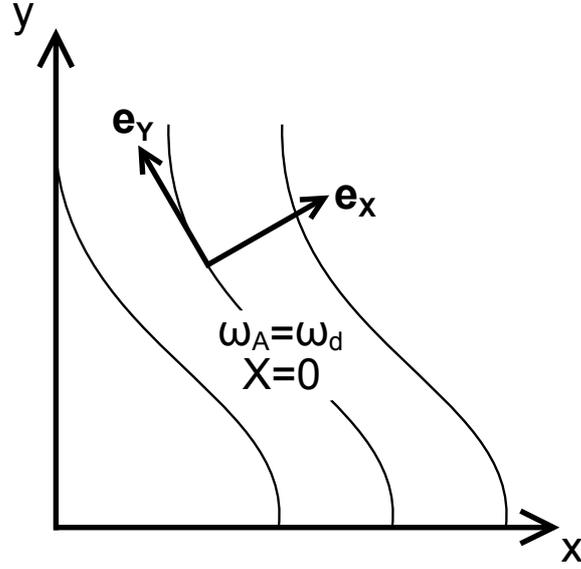


Figure 2.2: Contours of field-line eigenfrequency, and the two coordinate systems used in this chapter. In addition to using Cartesian coordinates  $(x, y, z)$  (in which a coordinate unit vector has fixed orientation everywhere in space) it is convenient to define a new system  $(X, Y, z)$  from contours of field-line eigenfrequency,  $\omega_A$ . In this new system,  $\hat{\mathbf{X}}$  is perpendicular to contours of  $\omega_A$  and  $\hat{\mathbf{Y}}$  is tangent to contours of  $\omega_A$ . If there is a contour at which  $\omega_A$  matches a driving frequency,  $\omega_d$ , then the origin of  $X$  is chosen so that  $\omega_A = \omega_d$  at  $X = 0$ .

### 2.4.1 Governing Equations

Neglecting gravity, the cold ( $\beta = 0$  or equivalently  $p = 0$ ), ideal ( $\eta = 0$  and  $\nu = 0$ ) MHD equations are

$$\text{Induction equation: } \frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}), \quad (2.4)$$

$$\text{Momentum equation: } \rho \frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u} \cdot \nabla) \mathbf{u} = \mathbf{j} \times \mathbf{B}, \quad (2.5)$$

$$\text{Continuity equation: } \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (2.6)$$

$$\text{Ampère's law: } \mathbf{j} = \frac{1}{\mu_0} \nabla \times \mathbf{B}, \quad (2.7)$$

$$\text{Solenoid constraint: } \nabla \cdot \mathbf{B} = 0. \quad (2.8)$$

Taking  $B_0 \hat{\mathbf{z}}$  as the equilibrium magnetic field and  $\rho_e$  as the equilibrium density, using Ampère's Law and neglecting high-order (non-linear) perturbations, the induction and momentum equations reduce to

$$\frac{\partial \mathbf{B}}{\partial t} = B_0 \left( \frac{\partial u_x}{\partial z}, \frac{\partial u_y}{\partial z}, -\frac{\partial u_x}{\partial x} - \frac{\partial u_y}{\partial y} \right), \quad (2.9)$$

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{B_0}{\mu_0 \rho_e} \left( \frac{\partial b_x}{\partial z} - \frac{\partial b_z}{\partial x}, \frac{\partial b_y}{\partial z} - \frac{\partial b_z}{\partial y}, 0 \right). \quad (2.10)$$

As matters stand  $u_x$ ,  $u_y$ ,  $b_x$ ,  $b_y$  and  $b_z$  are functions of  $x$ ,  $y$ ,  $z$  and  $t$ . We consider perfectly reflecting boundaries in the  $z$ -direction (equivalent to a perfectly conducting ionosphere in a magnetospheric model), allowing us to take Fourier modes in the  $z$  direction. Writing

$$\begin{aligned} b_x(x, y, z, t) &\rightarrow b_x(x, y, t) \cos(k_z z), \\ b_y(x, y, z, t) &\rightarrow b_y(x, y, t) \cos(k_z z), \\ b_z(x, y, z, t) &\rightarrow b_z(x, y, t) \sin(k_z z), \\ u_x(x, y, z, t) &\rightarrow u_x(x, y, t) \sin(k_z z), \\ u_y(x, y, z, t) &\rightarrow u_y(x, y, t) \sin(k_z z), \end{aligned}$$

the problem reduces to 2D with governing equations

$$\frac{\partial b_x(x, y, t)}{\partial t} = B_0 u_x(x, y, t) k_z, \quad (2.11)$$

$$\frac{\partial b_y(x, y, t)}{\partial t} = B_0 u_y(x, y, t) k_z, \quad (2.12)$$

$$\frac{\partial b_z(x, y, t)}{\partial t} = -B_0 \left( \frac{\partial u_x(x, y, t)}{\partial x} + \frac{\partial u_y(x, y, t)}{\partial y} \right), \quad (2.13)$$

$$\frac{\partial u_x(x, y, t)}{\partial t} = -\frac{B_0}{\mu_0 \rho_e(x, y)} \left( k_z b_x(x, y, t) + \frac{\partial b_z(x, y, t)}{\partial x} \right), \quad (2.14)$$

$$\frac{\partial u_y(x, y, t)}{\partial t} = -\frac{B_0}{\mu_0 \rho_e(x, y)} \left( k_z b_y(x, y, t) + \frac{\partial b_z(x, y, t)}{\partial y} \right). \quad (2.15)$$

In preparation for numerical solution of the governing equations, we non-dimensionalise, setting,

$$\tilde{\mathbf{B}} = \frac{\mathbf{B}}{B_0}, \quad (2.16)$$

$$\tilde{\mathbf{u}} = \frac{\mathbf{u}}{v_{A0}}, \quad (2.17)$$

$$\tilde{\rho}_e = \frac{\rho_e}{\rho_0}, \quad (2.18)$$

$$\tilde{\mathbf{r}} = \frac{\mathbf{r}}{l_0}, \quad (2.19)$$

$$\tilde{t} = \frac{t}{\tau} = \frac{v_{A0} t}{l_0}, \quad (2.20)$$

where  $v_{A0}$  is the equilibrium Alfvén speed at  $(\tilde{x}, \tilde{y}) = (0, 0)$ ,  $\rho_0$  is the equilibrium density at  $(\tilde{x}, \tilde{y}) = (0, 0)$ , and  $l_0$  is the width of the rectangular domain in the  $x$ -direction. Since the system is linear, we are free to further scale perturbations by a dimensionless ‘smallness’ parameter.

Putting equations (2.11) to (2.15) into dimensionless form, we get,

$$\frac{\partial \tilde{b}_x(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{t}} = \tilde{u}_x(\tilde{x}, \tilde{y}, \tilde{t}) \tilde{k}_z, \quad (2.21)$$

$$\frac{\partial \tilde{b}_y(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{t}} = \tilde{u}_y(\tilde{x}, \tilde{y}, \tilde{t}) \tilde{k}_z, \quad (2.22)$$

$$\frac{\partial \tilde{b}_z(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{t}} = - \left( \frac{\partial \tilde{u}_x(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{x}} + \frac{\partial \tilde{u}_y(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{y}} \right), \quad (2.23)$$

$$\frac{\partial \tilde{u}_x(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{t}} = - \frac{1}{\tilde{\rho}_e(\tilde{x}, \tilde{y})} \left( \tilde{k}_z \tilde{b}_x(\tilde{x}, \tilde{y}, \tilde{t}) + \frac{\partial \tilde{b}_z(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{x}} \right), \quad (2.24)$$

$$\frac{\partial \tilde{u}_y(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{t}} = - \frac{1}{\tilde{\rho}_e(\tilde{x}, \tilde{y})} \left( \tilde{k}_z \tilde{b}_y(\tilde{x}, \tilde{y}, \tilde{t}) + \frac{\partial \tilde{b}_z(\tilde{x}, \tilde{y}, \tilde{t})}{\partial \tilde{y}} \right). \quad (2.25)$$

## 2.4.2 Method

Equations (2.21) to (2.25) were solved numerically on a finite domain. This section describes the treatment of boundary-conditions, the profile of Alfvén speed, driving, the numerical scheme, and tests that establish results are accurate.

### 2.4.2.1 Choice of Boundary-Conditions

Whilst the physics of field-line resonance is common to all MHD plasmas, boundary-conditions were chosen with Earth's magnetosphere in mind. The simulation is easier if the 'hockey stick' shape of the dusk flank is replaced with a rectangle as illustrated in Figure 2.3. This simplification loses the curved geometry near the nose of the magnetosphere, but the properties of the different boundaries are preserved: the antisunward boundary at large  $\tilde{y}$  is open to model the tail and reflections from this boundary were avoided by advancing it ahead of all perturbations; the boundary-condition at  $\tilde{y} = 0$  is symmetric, representing the nose of the magnetosphere; most fast waves propagating towards Earth are refracted and redirected towards the magnetopause, so a reflecting boundary was placed at  $\tilde{x} = 0$  (the solution should be fairly insensitive to this boundary-condition as fast waves are evanescent here); and finally, the boundary at  $\tilde{x} = 1$  was chosen as the magnetopause and may be reflecting or driven.

### 2.4.2.2 Profile of Alfvén Speed

The density profile was chosen so that  $\tilde{v}_A(\tilde{x}, \tilde{y})$  provides a 2D variation of field-line eigenfrequencies. In choosing  $\tilde{v}_A(\tilde{x}, \tilde{y})$  it is helpful to choose a function that is continuous and smooth,

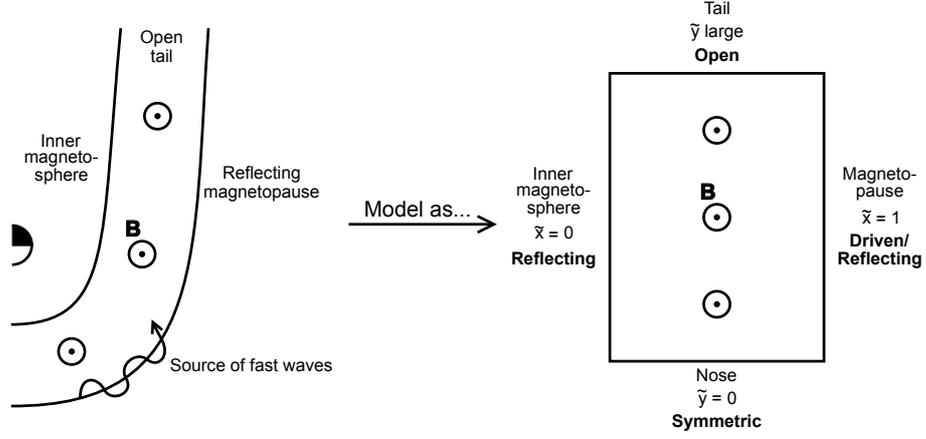


Figure 2.3: Equatorial plane cut of the magnetosphere, and the corresponding simulation domain. Boundary-conditions are chosen to mimic the flank of the magnetosphere, which acts as a waveguide for fast waves. Credit: Russell and Wright (2010), modified with permission © ESO.

and that can be controlled using a small number of parameters. For consistency with the outer magnetosphere, it is also desirable that Alfvén speed decrease with  $\tilde{x}$ . With these aims in mind, the function chosen for  $\tilde{v}_A(\tilde{x}, \tilde{y})$  was

$$\tilde{v}_A(\tilde{x}, \tilde{y}) = \begin{cases} 1, & \tilde{x} \leq \tilde{x}_-(\tilde{y}), \\ 1 - (1 - \tilde{v}_1) \sin^2 \left( \frac{\pi(\tilde{x} - \tilde{x}_-(\tilde{y}))}{2\tilde{L}_x} \right), & \tilde{x}_-(\tilde{y}) < \tilde{x} < \tilde{x}_+(\tilde{y}), \\ \tilde{v}_1, & \tilde{x}_+(\tilde{y}) \leq \tilde{x}, \end{cases} \quad (2.26)$$

where

$$\tilde{x}_-(\tilde{y}) = \begin{cases} (1 - \tilde{L}_x), & \tilde{y} \leq \tilde{y}_0, \\ (1 - \tilde{L}_x) \cos^2 \left( \frac{\pi(\tilde{y} - \tilde{y}_0)}{2\tilde{L}_y} \right), & \tilde{y}_0 < \tilde{y} < \tilde{y}_0 + \tilde{L}_y, \\ 0, & \tilde{y}_0 + \tilde{L}_y \leq \tilde{y}, \end{cases} \quad (2.27)$$

and  $\tilde{x}_+(\tilde{y}) = \tilde{x}_-(\tilde{y}) + \tilde{L}_x$ . This produces two uniform regions of Alfvén speed, the interior region ( $\tilde{x} \leq \tilde{x}_-(\tilde{y})$ ) having  $\tilde{v}_A(\tilde{x}, \tilde{y}) = 1$  and the exterior region ( $\tilde{x} \geq \tilde{x}_+(\tilde{y})$ ) having  $\tilde{v}_A(\tilde{x}, \tilde{y}) = \tilde{v}_1$ . We are free to choose  $\tilde{v}_1$ , but require it to be less than 1 so that Alfvén speed reduces with increasing  $\tilde{x}$ . The uniform regions are joined by a corridor of width  $\tilde{L}_x$  in which the Alfvén speed is non-uniform. Figure 2.4 shows a contour plot of  $\tilde{v}_A(\tilde{x}, \tilde{y})$  for  $\tilde{v}_1 = 0.2$ ,  $\tilde{y}_0 = 1$ ,  $\tilde{L}_x = 0.4$  and  $\tilde{L}_y = 4$ . Figure 2.5 shows the variation of  $\tilde{v}_A(\tilde{x}, \tilde{y})$  on fixed  $\tilde{y}$  for these same parameters, making the cut where  $\tilde{x}_-(\tilde{y}) = 0.2$ .

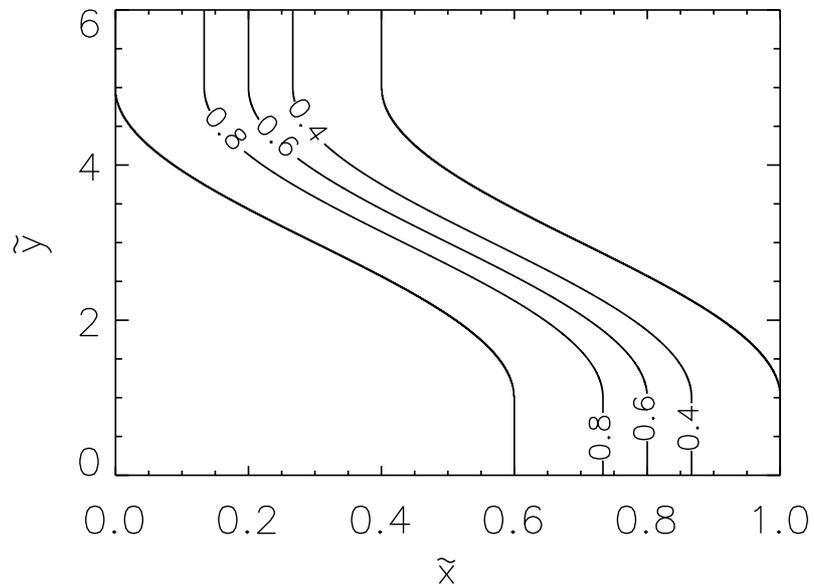


Figure 2.4: Contour plot of  $\tilde{v}_A(\tilde{x}, \tilde{y})$  for sample parameters. This plot was produced for  $\tilde{v}_1 = 0.2$ ,  $\tilde{y}_0 = 1$ ,  $\tilde{L}_x = 0.4$  and  $\tilde{L}_y = 4$ . The contour spacing is  $\Delta\tilde{v}_A = 0.2$ .

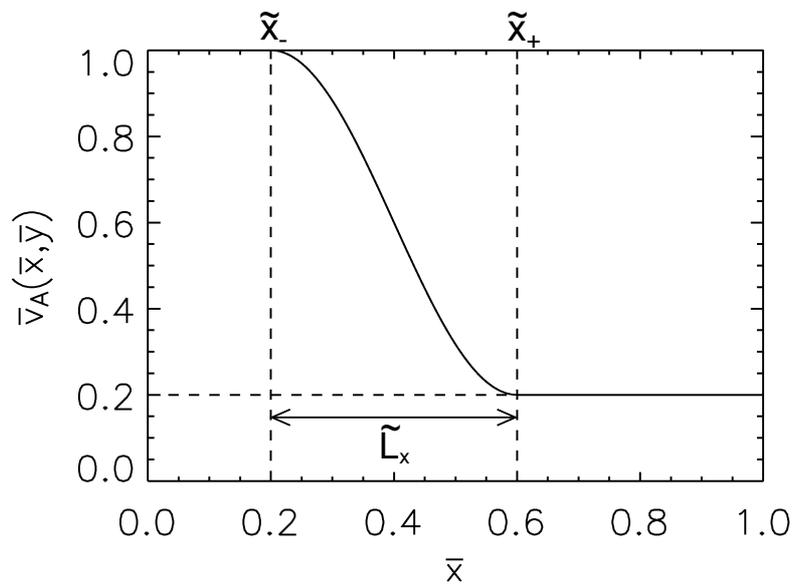


Figure 2.5: Cut in  $\tilde{v}_A(\tilde{x}, \tilde{y})$  on fixed  $\tilde{y}$ . This plot was produced taking  $\tilde{v}_1 = 0.2$ ,  $\tilde{L}_x = 0.4$  and  $\tilde{x}_-(\tilde{y}) = 0.2$ . Credit: Russell and Wright (2010), modified with permission © ESO.

### 2.4.2.3 Driving

The fast wave was driven by setting  $\tilde{u}_x$  at  $\tilde{x} = 1$ . At this boundary, the displacement in the  $\tilde{x}$  direction is a sinusoidal function ramped in  $\tilde{y}$  and  $\tilde{t}$ . This function is continuous and differentiable in  $\tilde{y}$  and  $\tilde{t}$ , and produces no net (time averaged) displacement of the boundary. Initially, the amplitude of the displacement ramps up globally over  $N_t$  periods; under this envelope, the displacement is a sinusoidal wave that ramps up spatially over one wavelength, is at full amplitude for  $N_y$  wavelengths and ramps down over one wavelength. Differentiating the displacement in time to obtain a velocity, we set

$$\tilde{u}_x(\tilde{y}, \tilde{t}) = \begin{cases} f(\tilde{y}) \left( \frac{1}{4N_t} \sin\left(\frac{\pi\tilde{t}}{\tilde{t}_1}\right) \sin(\tilde{k}_d\tilde{y} - \tilde{\omega}_d\tilde{t}) - \sin^2\left(\frac{\pi\tilde{t}}{2\tilde{t}_1}\right) \cos(\tilde{k}_d\tilde{y} - \tilde{\omega}_d\tilde{t}) \right), & \tilde{t} < \tilde{t}_1, \\ -f(\tilde{y}) \cos(\tilde{k}_d\tilde{y} - \tilde{\omega}_d\tilde{t}), & \tilde{t} \geq \tilde{t}_1, \end{cases} \quad (2.28)$$

where  $\tilde{t}_1 = 2\pi N_t / \tilde{\omega}_d$  and

$$f(\tilde{y}) = \begin{cases} \frac{1}{2} \left( 1 - \cos\left(\frac{\tilde{k}_d\tilde{y}}{2}\right) \right), & \tilde{y} < \tilde{y}_1 \\ 1, & \tilde{y}_1 \leq \tilde{y} \leq \tilde{y}_2 \\ \frac{1}{2} \left( 1 + \cos\left(\frac{\tilde{k}_d}{2}(\tilde{y} - \tilde{y}_2)\right) \right), & \tilde{y}_2 < \tilde{y} < \tilde{y}_3, \\ 0, & \tilde{y}_3 \leq \tilde{y}, \end{cases} \quad (2.29)$$

with  $\tilde{y}_1 = 2\pi/\tilde{k}_d$ ,  $\tilde{y}_2 = 2\pi(N_y + 1)/\tilde{k}_d$  and  $\tilde{y}_3 = 2\pi(N_y + 2)/\tilde{k}_d$ . An illustration of the driver, once it has fully ramped up in time, is provided in Figure 2.6.

We chose a continuous monochromatic driver because it shows the physics of field-line resonance clearly and cleanly. It can also be viewed as a building block for systems driven at multiple frequencies, because a non-uniform system supports collective modes of oscillation with discrete eigenfrequencies. When such a system is driven by a broadband source that includes one or more of its eigenfrequencies, interference leads to the dominance of those eigenfrequencies. Thus, a broadband driver drives resonances as if it were a superposition of monochromatic drivers (Kivelson and Southwood, 1985; Rickard and Wright, 1994; de Groof et al., 1998).

### 2.4.2.4 Numerical Scheme and Testing

Governing equations (2.21) to (2.25) were solved using the leapfrog-trapezoidal scheme detailed in Rickard and Wright (1994). Using centred differences for spatial derivatives, this code is second order accurate in space and time. Runs were performed with a uniform grid-spacing  $\Delta\tilde{l} \leq 0.00332$  and a time-step  $\Delta\tilde{t} = 0.8\Delta\tilde{l}$ . For each run, the grid-spacing was chosen to ensure the short spatial scales produced by phase-mixing were resolved with at least five points.

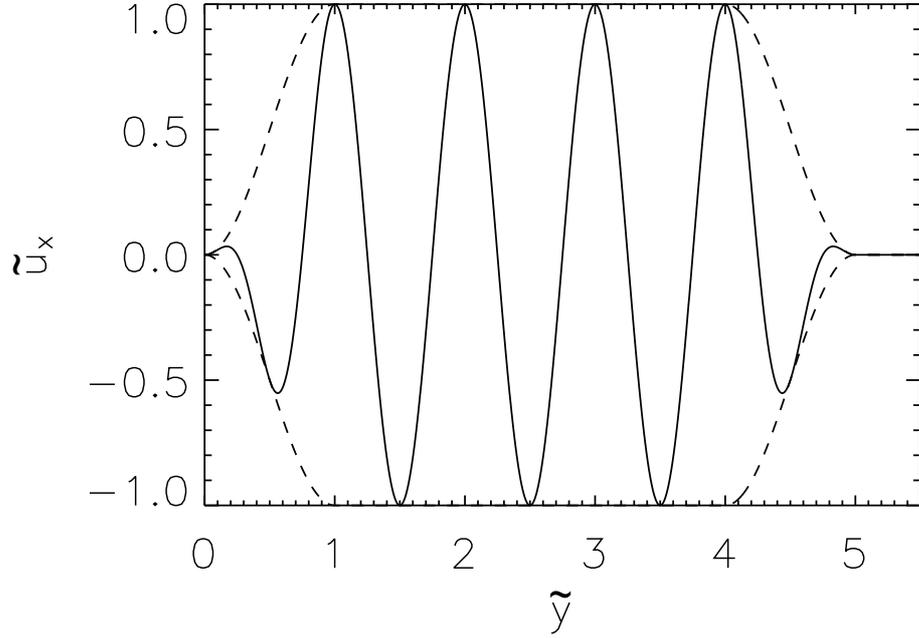


Figure 2.6: Illustration of the driver imposed at  $\tilde{x} = 1$ , with  $\tilde{k}_d = 2\pi$  and  $N_y = 3$ . The solid line shows a snapshot of  $\tilde{u}_x$ . Dashed lines show the envelope  $\pm f(\tilde{y})$ .

Four tests confirm that the code produces accurate results. First, the code conserves energy: total energy-density in the waveguide agrees with time-integrated Poynting flux across the boundaries, to within 0.0447% after early times (initially, total energy-density is zero, making a percentage measure unreliable). Second, the code maintains a small  $\tilde{\nabla} \cdot \tilde{\mathbf{B}}$ : because we do not damp  $\tilde{\nabla} \cdot \tilde{\mathbf{B}}$ , round-off errors do accumulate; nonetheless the maximum value in the runs shown here was  $1.65 \times 10^{-10}$ , so these errors are negligible. Test cases showed that the code captures refraction of wavefronts and their reflection at boundaries in  $\tilde{x}$ , and, as a final test, running the code with the drivers and density profile of Rickard and Wright (1994) reproduced their results.

## 2.4.3 Results

### 2.4.3.1 Excitation of Resonant Alfvén Wave

As a simulation runs, energy propagates throughout the domain. We are considering a cold plasma, so features of energy-density that move across magnetic field-lines correspond to the fast wave. This fast wave soon reaches a quasi-steady-state in the vicinity of the driven boundary, in which energy losses to the resonant Alfvén wave and flux down the waveguide approximately match energy fed in from the driven boundary.

The profile of Alfvén speed means that field-lines in our domain have eigenfrequencies that lie

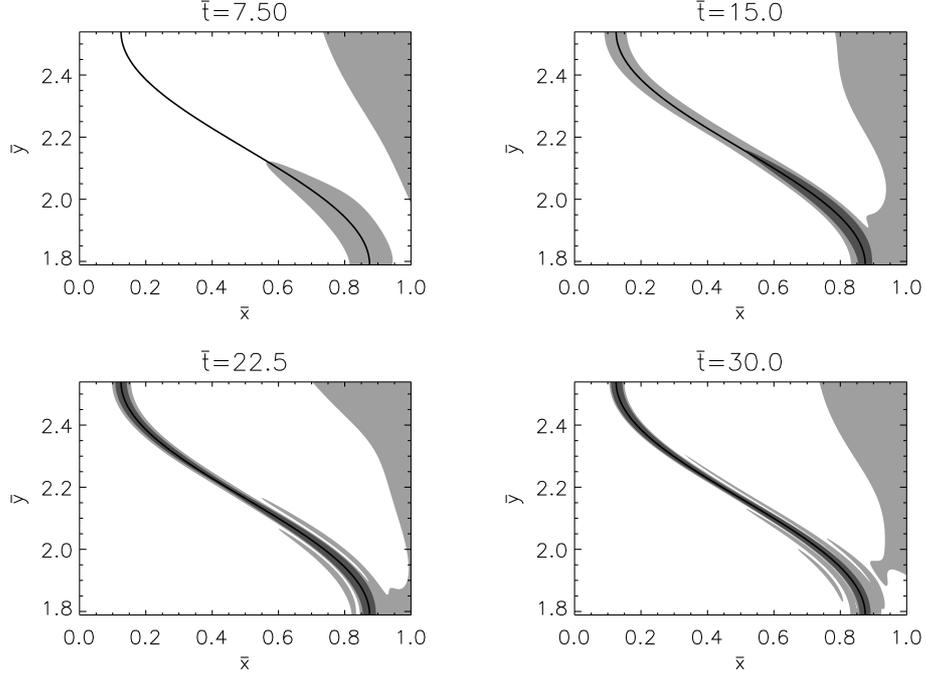


Figure 2.7: Accumulation of energy-density at a fully 2D resonant contour. Plots show filled contours of energy-density with contours representing 0.3% and 1.4% of the maximum energy-density at  $\tilde{t} = 30.0$ . The location of the resonant contour, where  $\tilde{\omega}_A(\tilde{x}, \tilde{y}) = \tilde{\omega}_d$ , is shown by the black line. Credit: Russell and Wright (2010), reproduced with permission © ESO.

in a continuum  $\tilde{\omega}_A(\tilde{x}, \tilde{y}) \in [\tilde{v}_1 \tilde{k}_z, \tilde{k}_z]$ . When the driving frequency,  $\tilde{\omega}_d$ , falls outside this continuum there is no (singular) field-line resonance and energy remains in the fast wave. However, when the driving frequency somewhere matches a field-line eigenfrequency, energy is deposited in the vicinity of the surface at which  $\tilde{\omega}_A(\tilde{x}, \tilde{y}) = \tilde{\omega}_d$ . Figure 2.7 shows this deposition of energy. The energy-density of the resonant Alfvén wave grows to more than 100 times the energy-density of the fast wave during the simulation.

At the resonance, the dominant velocity perturbation is tangential to the resonant surface. Furthermore, once deposited on a field-line, energy remains trapped there. For the case of decoupled modes, these properties are uniquely possessed by the Alfvén wave, so it is sensible to describe the deposited energy as belonging to a resonant Alfvén wave.

### 2.4.3.2 Phase-Mixing

Field-line resonance produces nearly perfect conditions for phase-mixing, because Alfvén waves are excited on a gradient of field-line Alfvén eigenfrequency. Figure 2.8 shows phase-mixing at work in our simulations, producing short length-scales in the resonant Alfvén wave, even as field-line resonance continues to deposit energy. The four panels show snapshots of  $u_Y$  (the velocity

perturbation tangent to the resonant surface) along a curve perpendicular to the resonant surface (i.e. on fixed  $Y$ ). At early times, the resonance is broad and gradients of the wave fields are gentle. As time progresses, however, the amplitude of the Alfvén wave increases as more energy is deposited by resonance (note the changing vertical scale in Figure 2.8) and length-scales shorten through phase-mixing. In this way, the perturbations that make up the Alfvén wave increase in amplitude, and the resonance becomes increasingly narrow, developing steep gradients.

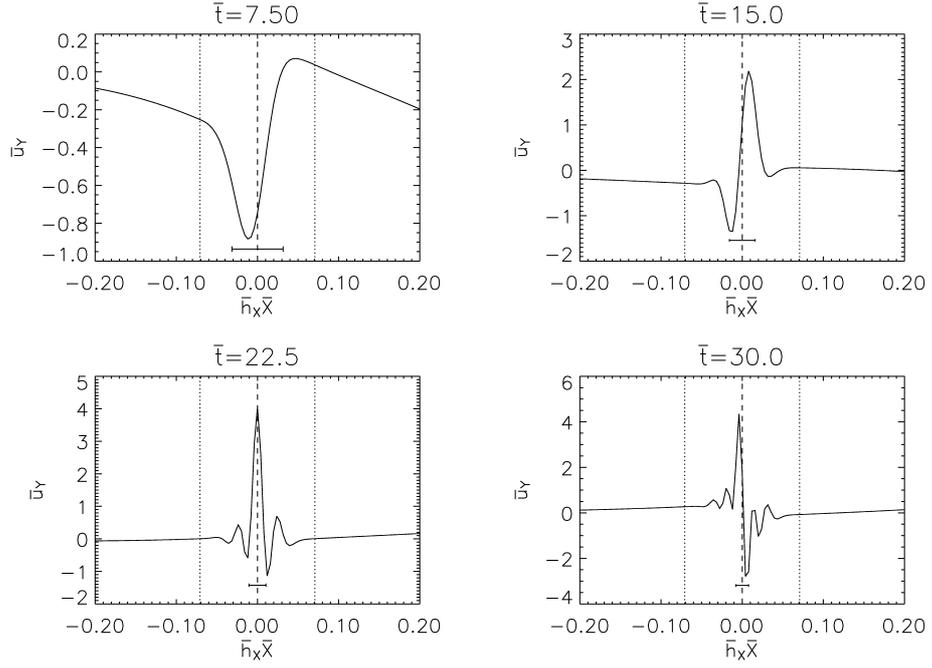


Figure 2.8: Phase-mixing of velocity component tangent to the resonant contour. The plot is made along a curve that is everywhere perpendicular to contours of  $\tilde{\omega}_A(\tilde{x}, \tilde{y})$  (i.e. on fixed  $Y$ ). Dotted lines indicate the transition from uniform to non-uniform regions, with the dashed line showing the location at which  $\tilde{\omega}_A(\tilde{x}, \tilde{y}) = \tilde{\omega}_d$ . The horizontal bar indicates the dimensionless phase-mixing length which is proportional to  $1/\tilde{t}$ . Note that the vertical axis changes between these plots, the velocity perturbation increasing in time. Credit: Russell and Wright (2010), reproduced with permission © ESO.

For a similar 1D model in which  $v_A = v_A(x)$ , Mann et al. (1995) (see also the analysis in Section 1.2.5) showed that the shortest length-scale within a phase-mixing Alfvén wave is governed by a time-dependent phase-mixing length,

$$L_{\text{ph}}(t) = 2\pi \left( t \frac{d\omega_A}{dx} \right)^{-1}. \quad (2.30)$$

As the resonance becomes increasingly narrow, the wavelength perpendicular to the resonant surface becomes much less than both the wavelength along the resonant surface and the radius of curvature for the resonant surface. In this limit, we expect the 1D phase-mixing length to provide an increasingly good description of the resonant Alfvén wave, provided the coordinate is suitably

handled. By this reasoning, we surmise that this result might be extended to higher dimensions by taking

$$L_{\text{ph}}(t) = 2\pi (t|\nabla\omega_A|)^{-1}, \quad (2.31)$$

and this is in good agreement with the fine scales observed in the simulations (see the horizontal line segments indicated in Figure 2.8).

### 2.4.3.3 Imprinting of Spatial Forms

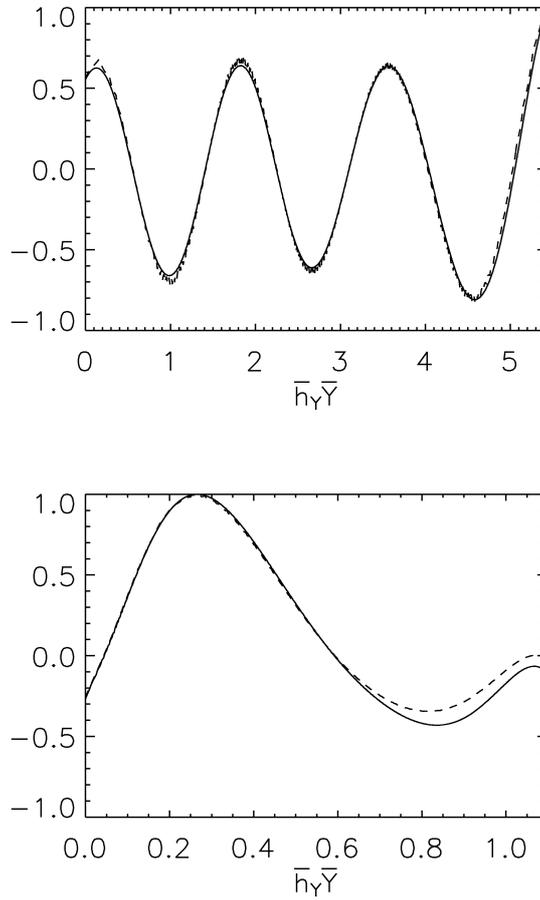


Figure 2.9: Plots at  $\tilde{t} = 30$  of Alfvénic velocity component,  $\tilde{u}_Y$ , (dashed curve) and magnetic-pressure gradient,  $\sim -\partial\tilde{b}_z/\partial\tilde{s}$ , (solid curve) against distance,  $\tilde{s} = \int \bar{h}_Y d\bar{Y}$ , along the resonant contour. These quantities have been re-normalised by their maximum value to aid comparison of spatial forms. Upper plot: a quasi-1D case. Lower plot: a fully 2D case. Credit: Russell and Wright (2010), reproduced with permission © ESO.

At any time, the resonant Alfvén wave varies significantly along the resonant contour and the

associated velocity perturbation,  $\tilde{u}_Y$ , may change sign. This is in contrast to decoupled Alfvén waves, which are invariant along contours of field-line Alfvén eigenfrequency. In Section 2.5, an analytic solution is obtained that describes the limit  $t \rightarrow \infty$ . A key result is that the dominant velocity perturbation is parallel to the resonant surface and given by equation (2.175):

$$u_Y = -i \left( \frac{B_0}{\mu_0 \omega [\partial \rho / \partial X]_{X=0}} \right) \left( \frac{1}{h_{Y0}} \left[ \frac{\partial b_z}{\partial Y} \right]_{X=0} \right) X^{-1}.$$

The analytic result is singular at the resonant surface, but the result reveals a striking dependence on the magnetic-pressure gradient associated with the fast wave  $\sim -\partial \tilde{b}_z / \partial \tilde{s}$  where  $d\tilde{s} = \tilde{h}_{Y0} d\tilde{Y}$  is the elemental distance along the resonant contour.

Our simulations show that the above relationship is also present in time-dependent problems, with strong correlation between spatial variations of the resonant Alfvén wave and the spatial form of the fast wave (Figure 2.9). This relationship is reasonable, because the velocity perturbation of the resonant Alfvén wave is driven by the magnetic-pressure gradient of the fast wave.

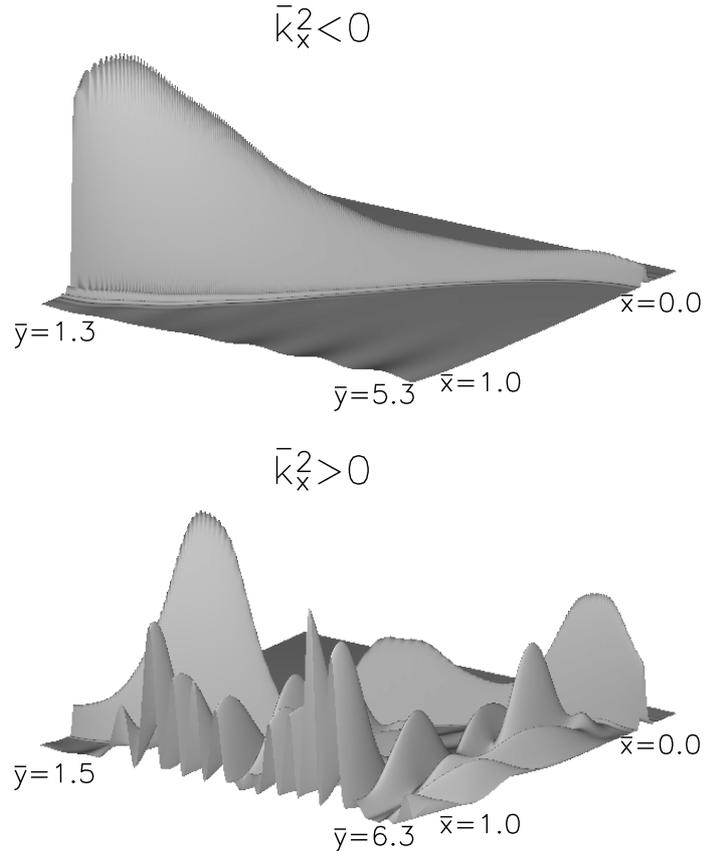


Figure 2.10: Surface plots of total energy-density at  $\tilde{t} = 30$  showing qualitative changes with  $\tilde{k}_x^2$ . Credit: Russell and Wright (2010), reproduced with permission © ESO.

We further investigated cases in which field-line eigenfrequencies were quasi-1D ( $\tilde{\omega}_A(\tilde{x}, \tilde{y})$ )

varies little with  $\tilde{y}$  over the wavelength of the driver). In this limit, the spatial form of the fast wave on the resonant contour is most strongly determined by the wavenumber  $\tilde{k}_x^2$  in the uniform region next to the driven boundary, which we can determine since

$$\tilde{k}_x^2 = \frac{\tilde{\omega}_d^2}{v_1^2} - \tilde{k}_z^2 - \tilde{k}_d^2, \quad (2.32)$$

for  $\tilde{x} > \tilde{x}_+$ . The driving frequency,  $\tilde{\omega}_d$ , which appears in the above equation, was set to match the Alfvén frequency at  $\tilde{x} = \tilde{x}_- + \tilde{L}_x/2$ . Selecting  $\tilde{k}_d$  such that  $\tilde{k}_x^2 < 0$  for  $\tilde{x} > \tilde{x}_+$ , gives a fast wave that is evanescent in  $\tilde{x}$ , over the entire numerical domain. If  $\tilde{k}_x^2 > 0$  for  $\tilde{x} > \tilde{x}_+$ , then the fast wave has a propagating character in  $\tilde{x}$ , between  $\tilde{x} = 1$  and the turning point of the fast wave.

For these cases, the relationship between the spatial form of the fast wave and the amplitude of the resonant Alfvén wave can be seen in surface plots of total energy-density. Figure 2.10 shows two such plots at  $\tilde{t} = 30$ . In each plot, the largest values of total energy-density lie on the resonant contour and are associated with the resonant Alfvén wave. Both plots show a foreground of fast wave energy, which lies between the resonance and the driven boundary. In the snapshot for  $\tilde{k}_x^2 < 0$ , the Alfvén wave has been driven to sufficiently large amplitude that the foreground appears almost negligible. In the snapshot for  $\tilde{k}_x^2 > 0$ , the foreground is much more visible; here the Alfvén wave corresponds to the triple peaked surface behind the fast wave foreground.

When  $\tilde{k}_x^2 < 0$  everywhere, there is less fast wave energy available to drive the resonance far from the driven boundary than there is close to the driven boundary, and this is apparent in the energy-density of the Alfvén wave. Setting  $\tilde{k}_x^2 > 0$  for  $\tilde{x} > \tilde{x}_+$ , means that (after initial transients) the fast wave forms an interference pattern, which may include nodes and anti-nodes. These nodes and anti-nodes prescribe points along the resonant contour at which energy is not available to the resonance or is available in maximum quantity. This, in turn, leads to the formation of nodes and anti-nodes in the energy-density of the Alfvén wave.

## 2.5 Analytic Solution

The simulations presented in Section 2.4 have demonstrated that field-line resonance persists in 2D equilibria; having shown its persistence, we would now like to set 2D field-line resonance on a firm mathematical footing. With this aim, we seek an analytic solution for late times, assuming that perturbations settle towards a time-dependence of the form  $\exp(-i\omega t)$ .

### 2.5.1 Governing Equations

In Section 2.4.1, governing equations were written for cold, linear, ideal MHD, using Cartesian coordinates. The equivalent equations are now written for curvilinear coordinates  $(X, Y, z)$ , as introduced in Section 2.3. So doing, the induction and momentum equations become

$$\frac{\partial \mathbf{B}}{\partial t} = B_0 \left( \frac{\partial u_X}{\partial z}, \frac{\partial u_Y}{\partial z}, -\frac{1}{h_X h_Y} \left( \frac{\partial}{\partial X} (h_Y u_X) + \frac{\partial}{\partial Y} (h_X u_Y) \right) \right), \quad (2.33)$$

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{B_0}{\mu_0 \rho_e} \left( \left( \frac{\partial b_X}{\partial z} - \frac{1}{h_X} \frac{\partial b_z}{\partial X} \right), \left( \frac{\partial b_Y}{\partial z} - \frac{1}{h_Y} \frac{\partial b_z}{\partial Y} \right), 0 \right). \quad (2.34)$$

Assuming perturbations have a  $(z, t)$ -dependence of the form  $\exp(i(k_z z - \omega t))$  and writing  $u_X = \partial \xi_X / \partial t \equiv -i\omega \xi_X$ ,  $u_Y = \partial \xi_Y / \partial t \equiv -i\omega \xi_Y$ , equations (2.33) and 2.34) reduce to,

$$b_X = iB_0 k_z \xi_X, \quad (2.35)$$

$$b_Y = iB_0 k_z \xi_Y, \quad (2.36)$$

$$b_z = -\frac{B_0}{h_X h_Y} \left[ \frac{\partial}{\partial X} (h_Y \xi_X) + \frac{\partial}{\partial Y} (h_X \xi_Y) \right], \quad (2.37)$$

$$\mathcal{L} \xi_X = \frac{B_0}{h_X} \frac{\partial b_z}{\partial X}, \quad (2.38)$$

$$\mathcal{L} \xi_Y = \frac{B_0}{h_Y} \frac{\partial b_z}{\partial Y}, \quad (2.39)$$

where

$$\mathcal{L}(X) = -k_z^2 B_0^2 + \omega^2 \mu_0 \rho_e(X). \quad (2.40)$$

### 2.5.2 Strategy

The governing equations can be solved by the following strategy:

1. Eliminate  $\xi_X$  and  $\xi_Y$  to obtain a single equation for  $b_z$ .
2. Solve for  $b_z$  using a series method.
3. Solve for series solutions of  $\xi_X$  and  $\xi_Y$  from  $b_z$ .

Solutions for  $b_X$  and  $b_Y$  will not be written explicitly, since these are simply multiples of  $\xi_X$  and  $\xi_Y$  respectively (2.35, 2.36). Where necessary, resonant and non-resonant cases will be considered separately, although the same techniques are used.

### 2.5.3 Preparation for Series Solution

#### 2.5.3.1 Elimination in favour of $b_z$

We can obtain a single partial differential equation (PDE) for  $b_z$  by using equations (2.38) and (2.39) to eliminate  $\xi_X$  and  $\xi_Y$  from (2.37). From (2.38) it follows that

$$\begin{aligned}
\mathcal{L}h_Y\xi_X &= B_0\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X} \\
\Rightarrow \frac{\partial \mathcal{L}}{\partial X}h_Y\xi_X + \mathcal{L}\frac{\partial}{\partial X}(h_Y\xi_X) &= B_0\frac{\partial}{\partial X}\left(\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X}\right) \\
\Rightarrow \mathcal{L}\frac{\partial}{\partial X}(h_Y\xi_X) &= B_0\frac{\partial}{\partial X}\left(\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X}\right) - \frac{\partial \mathcal{L}}{\partial X}h_Y\xi_X \\
\Rightarrow \mathcal{L}^2\frac{\partial}{\partial X}(h_Y\xi_X) &= B_0\mathcal{L}\frac{\partial}{\partial X}\left(\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X}\right) - \frac{\partial \mathcal{L}}{\partial X}\mathcal{L}h_Y\xi_X \\
\Rightarrow \mathcal{L}^2\frac{\partial}{\partial X}(h_Y\xi_X) &= B_0\mathcal{L}\frac{\partial}{\partial X}\left(\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X}\right) - B_0\frac{h_Y}{h_X}\frac{\partial \mathcal{L}}{\partial X}\frac{\partial b_z}{\partial X}.
\end{aligned} \tag{2.41}$$

From (2.39) it follows that

$$\begin{aligned}
\mathcal{L}^2h_X\xi_Y &= B_0\mathcal{L}\frac{h_X}{h_Y}\frac{\partial b_z}{\partial Y} \\
\Rightarrow \mathcal{L}^2\frac{\partial}{\partial Y}(h_X\xi_Y) &= B_0\mathcal{L}\frac{\partial}{\partial Y}\left(\frac{h_X}{h_Y}\frac{\partial b_z}{\partial Y}\right).
\end{aligned} \tag{2.42}$$

Substituting these into (2.37),

$$\begin{aligned}
\mathcal{L}^2b_z &= -\frac{B_0}{h_Xh_Y}\left[\mathcal{L}^2\frac{\partial}{\partial X}(h_Y\xi_X) + \mathcal{L}^2\frac{\partial}{\partial Y}(h_X\xi_Y)\right] \\
\Rightarrow \mathcal{L}^2b_z &= -\frac{B_0}{h_Xh_Y}\left[B_0\mathcal{L}\frac{\partial}{\partial X}\left(\frac{h_Y}{h_X}\frac{\partial b_z}{\partial X}\right) - B_0\frac{h_Y}{h_X}\frac{\partial \mathcal{L}}{\partial X}\frac{\partial b_z}{\partial X} + B_0\mathcal{L}\frac{\partial}{\partial Y}\left(\frac{h_X}{h_Y}\frac{\partial b_z}{\partial Y}\right)\right] \\
\Rightarrow -\frac{1}{B_0^2}h_Xh_Y\mathcal{L}^2b_z &= -\frac{h_Y}{h_X}\frac{\partial \mathcal{L}}{\partial X}\frac{\partial b_z}{\partial X} + \frac{h_Y}{h_X}\mathcal{L}\frac{\partial^2 b_z}{\partial X^2} + \frac{h_X}{h_Y}\mathcal{L}\frac{\partial^2 b_z}{\partial Y^2} \\
&\quad + \mathcal{L}\frac{\partial b_z}{\partial X}\frac{1}{h_X^2}\left(h_X\frac{\partial h_Y}{\partial X} - h_Y\frac{\partial h_X}{\partial X}\right) \\
&\quad + \mathcal{L}\frac{\partial b_z}{\partial Y}\frac{1}{h_Y^2}\left(h_Y\frac{\partial h_X}{\partial Y} - h_X\frac{\partial h_Y}{\partial Y}\right) \\
\Rightarrow -\frac{1}{B_0^2}h_X^3h_Y^3\mathcal{L}^2b_z &= h_Y^3h_X\mathcal{L}\frac{\partial^2 b_z}{\partial X^2} + h_X^3h_Y\mathcal{L}\frac{\partial^2 b_z}{\partial Y^2} \\
&\quad + \left[\mathcal{L}h_Y^2\left(h_X\frac{\partial h_Y}{\partial X} - h_Y\frac{\partial h_X}{\partial X}\right) - h_Y^3h_X\frac{\partial \mathcal{L}}{\partial X}\right]\frac{\partial b_z}{\partial X} \\
&\quad + \mathcal{L}h_X^2\left(h_Y\frac{\partial h_X}{\partial Y} - h_X\frac{\partial h_Y}{\partial Y}\right)\frac{\partial b_z}{\partial Y}.
\end{aligned} \tag{2.43}$$

Equation (2.43) is a homogeneous PDE for  $b_z$  that is second order in  $X$ . As written above, it appears fairly substantial, so it is both reassuring and helpful to write it as

$$F \frac{\partial^2 b_z}{\partial X^2} + G \frac{\partial b_z}{\partial X} + \hat{H} b_z = 0, \quad (2.44)$$

where

$$F = h_Y^3 h_X \mathcal{L}, \quad (2.45)$$

$$G = \mathcal{L} h_Y^2 \left( h_X \frac{\partial h_Y}{\partial X} - h_Y \frac{\partial h_X}{\partial X} \right) - h_Y^3 h_X \frac{\partial \mathcal{L}}{\partial X}, \quad (2.46)$$

$$\hat{H} = \frac{1}{B_0^2} h_X^3 h_Y^3 \mathcal{L}^2 + \mathcal{L} h_X^2 \left( h_Y \frac{\partial h_X}{\partial Y} - h_X \frac{\partial h_Y}{\partial Y} \right) \frac{\partial}{\partial Y} + h_X^3 h_Y \mathcal{L} \frac{\partial^2}{\partial Y^2}. \quad (2.47)$$

All derivatives with respect to  $Y$  are encompassed in  $\hat{H}$ , which is written with a hat to emphasise that it operates on functions of  $Y$  to its right. Equation (2.44) is the single equation for  $b_z$  that must be solved in both resonant and non-resonant cases.

### 2.5.3.2 Expansions

The standard method for solving an ordinary differential equation (ODE) with variable coefficients and a singularity is the method of Frobenius (e.g. Bender and Orszag (1978)). In adapting this method to a 2D PDE we are guided by Thompson and Wright (1993) and seek a series solution for  $b_z$  of the form

$$b_z(X, Y) = X^\sigma \sum_{n=0}^{\infty} X^n \alpha_n(Y) + X^\sigma \ln(X) \sum_{n=0}^{\infty} X^n \beta_n(Y). \quad (2.48)$$

Note that

$$\frac{\partial b_z}{\partial X} = X^{\sigma-1} \sum_{n=0}^{\infty} X^n [(\sigma+n)\alpha_n + \beta_n] + X^{\sigma-1} \ln(X) \sum_{n=0}^{\infty} X^n (\sigma+n)\beta_n, \quad (2.49)$$

$$\begin{aligned} \frac{\partial^2 b_z}{\partial X^2} &= X^{\sigma-2} \sum_{n=0}^{\infty} X^n [(\sigma+n)(\sigma+n-1)\alpha_n + (2\sigma+2n-1)\beta_n] \\ &\quad + X^{\sigma-2} \ln(X) \sum_{n=0}^{\infty} X^n (\sigma+n)(\sigma+n-1)\beta_n. \end{aligned} \quad (2.50)$$

Here,  $\mathcal{L}$  and the scale factors will be required to be regular, and are expanded as Taylor series in  $X$  constant  $Y$ , i.e.

$$\mathcal{L}(X) = \sum_{n=0}^{\infty} X^n \mathcal{L}_n(Y), \quad (2.51)$$

where

$$\mathcal{L}_0 = -k_z^2 B_0^2 + \omega^2 \mu_0 \rho_e(0), \quad (2.52)$$

$$\mathcal{L}_1 = \omega^2 \mu_0 \left[ \frac{\partial \rho_e}{\partial X} \right]_{X=0}, \quad (2.53)$$

$$\mathcal{L}_n = \frac{\omega^2}{n!} \mu_0 \left[ \frac{\partial^n \rho_e}{\partial X^n} \right]_{X=0}, \quad (2.54)$$

and

$$h_X(X, Y) = \sum_{n=0}^{\infty} X^n h_{Xn}(Y), \quad (2.55)$$

$$h_Y(X, Y) = \sum_{n=0}^{\infty} X^n h_{Yn}(Y), \quad (2.56)$$

where

$$h_{Xn}(Y) = \frac{1}{n!} \left[ \frac{\partial^n h_X}{\partial X^n} \right]_{(X,Y)=(0,Y)}, \quad (2.57)$$

$$h_{Yn}(Y) = \frac{1}{n!} \left[ \frac{\partial^n h_Y}{\partial X^n} \right]_{(X,Y)=(0,Y)}. \quad (2.58)$$

In the region of interest,  $h_X \neq 0$  and  $h_Y \neq 0$ ; in particular, this gives  $h_{X0} \neq 0$  and  $h_{Y0} \neq 0$ .

$F$ ,  $G$  and  $\hat{H}$  may themselves be expanded as Taylor series with coefficients  $F_n$ ,  $G_n$  and  $\hat{H}_n$ . These coefficients are obtained in terms of  $h_{Xn}$ ,  $h_{Yn}$  and  $\mathcal{L}_n$  by substituting (2.51), (2.55) and (2.56) into equations (2.45), (2.46) and (2.47) then collecting terms in  $X$ . Using a prime to indicate that a quantity has been differentiated with respect to  $Y$ , the first few Taylor coefficients for  $F$ ,  $G$  and  $\hat{H}$  are,

$$F_0 = \mathcal{L}_0 h_{Y0}^3 h_{X0}, \quad (2.59)$$

$$F_1 = \mathcal{L}_0 h_{Y0}^2 (3h_{Y1} h_{X0} + h_{Y0} h_{X1}) + \mathcal{L}_1 h_{Y0}^3 h_{X0}, \quad (2.60)$$

$$G_0 = \mathcal{L}_0 h_{Y0}^2 (h_{X0} h_{Y1} - h_{Y0} h_{X1}) - \mathcal{L}_1 h_{Y0}^3 h_{X0}, \quad (2.61)$$

$$\begin{aligned} G_1 = & 2\mathcal{L}_0 h_{Y0} (h_{X0} h_{Y1}^2 + h_{X0} h_{Y0} h_{Y2} - h_{X1} h_{Y0} h_{Y1} - h_{X2} h_{Y0}) \\ & - 2\mathcal{L}_1 h_{Y0}^2 (h_{X0} h_{Y1} + h_{X1} h_{Y0}) \\ & - 2\mathcal{L}_2 h_{X0} h_{Y0}^3, \end{aligned} \quad (2.62)$$

$$\begin{aligned} \hat{H}_0 = & \frac{1}{B_0^2} \mathcal{L}_0^2 h_{X0}^3 h_{Y0}^3 \\ & + \mathcal{L}_0 h_{X0}^2 (h_{Y0} h'_{X0} - h_{X0} h'_{Y0}) \frac{\partial}{\partial Y} \\ & + \mathcal{L}_0 h_{X0}^3 h_{Y0} \frac{\partial^2}{\partial Y^2}, \end{aligned} \quad (2.63)$$

$$\begin{aligned}
\hat{H}_1 &= \frac{1}{B_0^2} h_{X0}^2 h_{Y0}^2 \mathcal{L}_0 (3\mathcal{L}_0 (h_{X0} h_{Y1} + h_{X1} h_{Y0}) + 2\mathcal{L}_1 h_{X0} h_{Y0}) \\
&+ h_{X0} \left[ \mathcal{L}_0 \begin{pmatrix} h_{X0} h_{Y0} h'_{X1} + h_{X0} h_{Y1} h'_{X0} - h_{X0}^2 h'_{Y1} \\ -3h_{X0} h_{X1} h'_{Y0} + 2h_{X1} h_{Y0} h'_{X0} \\ + \mathcal{L}_1 h_{X0} (h_{Y0} h'_{X0} - h_{X0} h'_{Y0}), \end{pmatrix} \right] \frac{\partial}{\partial Y} \\
&+ h_{X0}^2 (\mathcal{L}_0 (h_{X0} h_{Y1} + 3h_{X1} h_{Y0}) + \mathcal{L}_1 h_{X0} h_{Y0}) \frac{\partial^2}{\partial Y^2}. \tag{2.64}
\end{aligned}$$

### 2.5.3.3 Equations for Coefficients of $b_z$

In order to solve equation (2.44) we solve for coefficients in the series representation of  $b_z$  i.e. for  $\alpha_n$  and  $\beta_n$  in equation (2.48). Substituting the expansions of Section 2.5.3.2 into equation (2.48),

$$\begin{aligned}
&\left( \sum_{n=0}^{\infty} X^n F_n \right) \left( X^{\sigma-2} \sum_{n=0}^{\infty} X^n [(\sigma+n)(\sigma+n-1)\alpha_n + (2\sigma+2n-1)\beta_n] \right. \\
&\quad \left. + X^{\sigma-2} \ln(X) \sum_{n=0}^{\infty} X^n (\sigma+n)(\sigma+n-1)\beta_n \right) \\
&+ \left( \sum_{n=0}^{\infty} X^n G_n \right) \left( X^{\sigma-1} \sum_{n=0}^{\infty} X^n [(\sigma+n)\alpha_n + \beta_n] \right. \\
&\quad \left. + X^{\sigma-1} \ln(X) \sum_{n=0}^{\infty} X^n (\sigma+n)\beta_n \right) \\
&+ \left( \sum_{n=0}^{\infty} X^n \hat{H}_n \right) \left( X^{\sigma} \sum_{n=0}^{\infty} X^n \alpha_n + X^{\sigma} \ln(X) \sum_{n=0}^{\infty} X^n \beta_n \right) \\
&= 0 \tag{2.65} \\
\Rightarrow &X^{\sigma-2} F_0 [\sigma(\sigma-1)\alpha_0 + (2\sigma-1)\beta_0] \\
&+ X^{\sigma-2} \ln(X) F_0 \sigma(\sigma-1)\beta_0 \\
&+ X^{\sigma-1} \begin{bmatrix} F_0(\sigma(\sigma+1)\alpha_1 + (2\sigma+1)\beta_1) \\ + F_1(\sigma(\sigma-1)\alpha_0 + (2\sigma-1)\beta_0) \\ + G_0(\sigma\alpha_0 + \beta_0) \end{bmatrix} \\
&+ X^{\sigma-1} \ln(X) [F_0 \sigma(\sigma+1)\beta_1 + F_1 \sigma(\sigma-1)\beta_0 + G_0 \sigma \beta_0] \\
&+ \sum_{n=0}^{\infty} X^{\sigma+n} \begin{bmatrix} \sum_{s=0}^{n+2} F_{n-s+2} [(\sigma+s)(\sigma+s-1)\alpha_s + (2\sigma+2s-1)\beta_s] \\ + \sum_{s=0}^{n+1} G_{n-s+1} [(\sigma+s)\alpha_s + \beta_s] \\ + \sum_{s=0}^n \hat{H}_{n-s} \alpha_s \end{bmatrix} \\
&+ \sum_{n=0}^{\infty} X^{\sigma+n} \ln(X) \begin{bmatrix} \sum_{s=0}^{n+2} F_{n-s+2} (\sigma+s)(\sigma+s-1)\beta_s \\ + \sum_{s=0}^{n+1} G_{n-s+1} (\sigma+s)\beta_s \\ + \sum_{s=0}^n \hat{H}_{n-s} \beta_s \end{bmatrix} \\
&= 0. \tag{2.66}
\end{aligned}$$

From this, we produce equations by matching powers of  $X$  and powers of  $X$  multiplied by  $\ln(X)$ .

Terms in  $X^{\sigma-2} \ln(X)$  give:

$$F_0 \sigma(\sigma - 1) \beta_0 = 0. \quad (2.67)$$

Terms in  $X^{\sigma-2}$  give:

$$F_0(\sigma(\sigma - 1)\alpha_0 + (2\sigma - 1)\beta_0) = 0. \quad (2.68)$$

Terms in  $X^{\sigma-1} \ln(X)$  give:

$$F_0 \sigma(\sigma + 1) \beta_1 + F_1 \sigma(\sigma - 1) \beta_0 + G_0 \sigma \beta_0 = 0. \quad (2.69)$$

Terms in  $X^{\sigma-1}$  give:

$$\begin{aligned} & F_0(\sigma(\sigma + 1)\alpha_1 + (2\sigma + 1)\beta_1) \\ & + F_1(\sigma(\sigma - 1)\alpha_0 + (2\sigma - 1)\beta_0) \\ & + G_0(\sigma\alpha_0 + \beta_0) \\ & = 0. \end{aligned} \quad (2.70)$$

Terms in  $X^{\sigma+n} \ln(X)$ ,  $n \geq 0$  give:

$$\begin{aligned} & F_0(\sigma + n + 2)(\sigma + n + 1)\beta_{n+2} \\ & + F_1(\sigma + n + 1)(\sigma + n)\beta_{n+1} \\ & + G_0(\sigma + n + 1)\beta_{n+1} \\ & + \sum_{s=0}^n \left( F_{n-s+2}(\sigma + s)(\sigma + s - 1) + G_{n-s+1}(\sigma + s) + \hat{H}_{n-s} \right) \beta_s \\ & = 0. \end{aligned} \quad (2.71)$$

Terms in  $X^{\sigma+n}$ ,  $n \geq 0$  give:

$$\begin{aligned} & F_0((\sigma + n + 2)(\sigma + n + 1)\alpha_{n+2} + (2\sigma + 2n + 3)\beta_{n+2}) \\ & + F_1((\sigma + n + 1)(\sigma + n)\alpha_{n+1} + (2\sigma + 2n + 1)\beta_{n+1}) \\ & + G_0((\sigma + n + 1)\alpha_{n+1} + \beta_{n+1}) \\ & + \sum_{s=0}^n \left( F_{n-s+2}(\sigma + s)(\sigma + s - 1) + G_{n-s+1}(\sigma + s) + \hat{H}_{n-s} \right) \alpha_s \\ & + \sum_{s=0}^n (F_{n-s+2}(2\sigma + 2s - 1) + G_{n-s+1}) \beta_s \\ & = 0. \end{aligned} \quad (2.72)$$

We solve these equations in the following two sections, treating the resonant and non-resonant

cases separately.

### 2.5.4 Non-Resonant Solution

In the non-resonant case, the frequency of the system,  $\omega$ , nowhere matches the natural field-line Alfvén frequency,  $\omega_A = k_z v_A = k_z B_0 / \sqrt{\mu_0 \rho_e}$ . We are free to choose the origin of  $X$  as any surface of  $\omega_A(X)$  near which we wish examine the solution. Referring back to (2.52), non-resonance means that  $\mathcal{L}_0 = \mathcal{L}(0) \neq 0$ , and hence  $F_0 \neq 0$  by (2.59). Having noted this, we now proceed with the solution.

#### 2.5.4.1 Value of $\beta_0$

If  $\sigma \notin \{0, 1\}$  (i.e. if  $\sigma$  is equal to neither 0 nor 1) then (2.67) gives  $\beta_0 = 0$ . If  $\sigma \in \{0, 1\}$  (i.e. if  $\sigma = 0$  or  $\sigma = 1$ ) then (2.68) gives  $\beta_0 = 0$ . Therefore,  $\beta_0 = 0 \forall \sigma$  (for all  $\sigma$ ).

#### 2.5.4.2 Value of $\beta_1$

If  $\sigma \notin \{0, -1\}$  then, since  $\beta_0 = 0$ , (2.69) gives  $\beta_1 = 0$ . If  $\sigma = 0$  then (2.70) gives  $\beta_1 = 0$ . If  $\sigma = -1$  then (2.68) gives  $\alpha_0 = 0$  and  $\beta_1 = 0$  follows immediately from (2.70). Therefore,  $\beta_1 = 0 \forall \sigma$ .

#### 2.5.4.3 Values of $\beta_{n+2}$ , $n \geq 0$

Assume that  $\beta_j = 0, \forall j < n + 2$ . This is the case for  $n = 0$  by the above. Then, equation (2.71) gives  $(\sigma + n + 1)(\sigma + n + 2)\beta_{n+2} = 0$ . It will be shown that  $\sigma \geq 0$ , so  $\beta_{n+2} = 0$ . By mathematical induction, it follows that  $\beta_{n+2} = 0 \forall n \geq 0$ . Combined with the results of Sections 2.5.4.1 and 2.5.4.2 this gives  $\beta_j = 0 \forall j \geq 0$ ; i.e. there are no terms involving logarithms in the non-resonant solution.

#### 2.5.4.4 Choosing $\sigma$ for General Solution

Putting  $\beta_0 = 0$  into (2.68) gives  $\sigma(\sigma - 1)\alpha_0 = 0$ . We require  $\alpha_0 \neq 0$  for a non-trivial solution, which implies  $\sigma \in \{0, 1\}$ . Since the possible values for  $\sigma$  are separated by an integer, the lowest value ( $\sigma = 0$ ) corresponds to the 2-parameter general solution (solving with  $\sigma = 1$  produces a 1-parameter solution that is a special case of the general solution).

### 2.5.4.5 Recurrence Relation for $\alpha_m$ , $m \geq 2$

Putting  $\sigma = \beta_j = 0$  into equation (2.72) gives, for  $n \geq 0$ ,

$$F_0(n+2)(n+1)\alpha_{n+2} + (nF_1 + G_0)(n+1)\alpha_{n+1} + \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \alpha_s = 0$$

$$\Rightarrow \alpha_{n+2} = -\frac{1}{(n+2)(n+1)F_0} \left[ \frac{(n+1)(nF_1 + G_0)\alpha_{n+1}}{+ \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \alpha_s} \right]$$

$$\Rightarrow \alpha_m = -\frac{1}{m(m-1)F_0} \left[ \frac{(m-1)((m-2)F_1 + G_0)\alpha_{m-1}}{+ \sum_{s=0}^{m-2} \left( s(s-1)F_{m-s} + sG_{m-s-1} + \hat{H}_{m-s-2} \right) \alpha_s} \right]. \quad (2.73)$$

This determines all coefficients  $\alpha_m$ ,  $m \geq 2$ , in terms of  $\alpha_0(Y)$  and  $\alpha_1(Y)$  by recursion relation, completing the 2-parameter solution.

### 2.5.4.6 General Solution as Sum of 1-Parameter Solutions

It is common to present a general, multi-parameter solution to a differential equation as a sum of independent 1-parameter solutions. This is familiar from the study of ODEs. There are a few changes for the PDE solved by the method above, most of which arise because parameters are functions of  $Y$ , and recursion relations for higher coefficients involve operators acting on functions of  $Y$ .

The first 1-parameter solution is obtained from the general solution by taking  $\alpha_0(Y) \neq 0$  and  $\alpha_1(Y) = 0$ . Then, recurrence relation (2.73) implies that  $\alpha_m$ ,  $m \geq 0$ , can be written as the result of an operation on  $\alpha_0$ , i.e.

$$\alpha_m = \hat{A}_{0,m} \alpha_0. \quad (2.74)$$

In particular, we have

$$\hat{A}_{0,0} = 1, \quad (2.75)$$

$$\hat{A}_{0,1} = 0. \quad (2.76)$$

The second 1-parameter solution is obtained by taking  $\alpha_0(Y) = 0$  and  $\alpha_1(Y) \neq 0$ . This gives

$$\alpha_m = \hat{A}_{1,m} \alpha_1, \quad (2.77)$$

with

$$\hat{A}_{1,0} = 0, \quad (2.78)$$

$$\hat{A}_{1,1} = 1. \quad (2.79)$$

For both solutions, remaining operators are found by substituting

$$\alpha_m = \hat{A}_{i,m} \alpha_i \quad (2.80)$$

into equation (2.73) to obtain the recurrence relation

$$\hat{A}_{i,m} = -\frac{1}{m(m-1)F_0} \left[ + \sum_{s=0}^{m-2} \left( s(s-1)F_{m-s} + sG_{m-s-1} + \hat{H}_{m-s-2} \right) \hat{A}_{i,s} \right], \quad (2.81)$$

for  $m \geq 2$ . Using this to evaluate the first few operators,

$$\hat{A}_{0,2} = -\frac{\hat{H}_0}{2F_0}, \quad (2.82)$$

$$\hat{A}_{0,3} = \frac{F_1 + G_0}{6F_0^2} \hat{H}_0 - \frac{\hat{H}_1}{6F_0}, \quad (2.83)$$

$$\hat{A}_{1,2} = -\frac{G_0}{2F_0}, \quad (2.84)$$

$$\hat{A}_{1,3} = \frac{F_1 + G_0}{6F_0^2} G_0 - \frac{G_1 + \hat{H}_0}{6F_0}. \quad (2.85)$$

From this, we may write the solution to the non-resonant problem as

$$b_z = \sum_{n=0}^{\infty} X^n \alpha_n, \quad (2.86)$$

$$b_z = \left( \sum_{n=0}^{\infty} X^n \hat{A}_{0,n} \right) \alpha_0(Y) + \left( \sum_{n=0}^{\infty} X^n \hat{A}_{1,n} \right) \alpha_1(Y), \quad (2.87)$$

$$b_z = \left[ 1 + X^2 \left( -\frac{\hat{H}_0}{2F_0} \right) + X^3 \left( \frac{F_1 + G_0}{6F_0^2} \hat{H}_0 - \frac{\hat{H}_1}{6F_0} \right) + \dots \right] \alpha_0(Y) \\ + \left[ X + X^2 \left( -\frac{G_0}{2F_0} \right) + X^3 \left( \frac{F_1 + G_0}{6F_0^2} G_0 - \frac{G_1 + \hat{H}_0}{6F_0} \right) + \dots \right] \alpha_1(Y). \quad (2.88)$$

2.5.4.7  $\xi_X$ 

Equation (2.38) states

$$\mathcal{L}h_X \xi_X = B_0 \frac{\partial b_z}{\partial X}.$$

In the non-resonant case,

$$\frac{\partial b_z}{\partial X} = \left( \sum_{n=0}^{\infty} n X^{n-1} \hat{A}_{0,n} \right) \alpha_0(Y) + \left( \sum_{n=0}^{\infty} n X^{n-1} \hat{A}_{1,n} \right) \alpha_1(Y). \quad (2.89)$$

Led by the form of  $\partial b_z / \partial X$  and equation (2.38), we let  $\xi_X$  take the form,

$$\xi_X = X^\sigma \frac{B_0}{\mathcal{L}_0 h_{X0}} \left( \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} \right) \alpha_0(Y) + X^\sigma \frac{B_0}{\mathcal{L}_0 h_{X0}} \left( \sum_{n=0}^{\infty} X^n \hat{C}_{1,n} \right) \alpha_1(Y), \quad (2.90)$$

where  $\sigma$  is once again a free parameter. Since  $b_z$  does not involve any logarithms we do not include them in our expansion of  $\xi_X$ .

Substituting in equation (2.38),

$$\begin{aligned} & \left( \sum_{n=0}^{\infty} X^n \mathcal{L}_n \right) \left( \sum_{n=0}^{\infty} X^n h_{Xn} \right) \frac{X^\sigma B_0}{\mathcal{L}_0 h_{X0}} \times \\ & \quad \left( \left( \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} \right) \alpha_0(Y) + \left( \sum_{n=0}^{\infty} X^n \hat{C}_{1,n} \right) \alpha_1(Y) \right) \\ & = B_0 \left( \sum_{n=0}^{\infty} n X^{n-1} \hat{A}_{0,n} \right) \alpha_0(Y) + B_0 \left( \sum_{n=0}^{\infty} n X^{n-1} \hat{A}_{1,n} \right) \alpha_1(Y). \end{aligned} \quad (2.91)$$

The lowest order, non-zero term on the right-hand side (RHS) is

$$B_0 \alpha_1(Y) X^0,$$

which must match with the lowest order, non-zero term on the left-hand side (LHS). Requiring that one of  $\hat{C}_{0,0}$  and  $\hat{C}_{1,0}$  be non-zero, the lowest order term on the LHS is

$$B_0 \left( \hat{C}_{0,0} \alpha_0 + \hat{C}_{1,0} \alpha_1 \right) X^\sigma,$$

so matching the LHS and RHS gives  $\sigma = 0$  and

$$\hat{C}_{0,0} = 0, \quad (2.92)$$

$$\hat{C}_{1,0} = 1. \quad (2.93)$$

Remaining operators are now obtained by matching terms in  $X^r \alpha_i$ . This gives

$$\begin{aligned}
& \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{i,s} = \mathcal{L}_0 h_{X0} (r+1) \hat{A}_{i,r+1} \\
\Rightarrow & \mathcal{L}_0 h_{X0} \hat{C}_{i,r} + \sum_{s=0}^{r-1} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{i,s} = \mathcal{L}_0 h_{X0} (r+1) \hat{A}_{i,r+1} \\
\Rightarrow & \hat{C}_{i,r} = (r+1) \hat{A}_{i,r+1} - \frac{1}{\mathcal{L}_0 h_{X0}} \sum_{s=0}^{r-1} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{i,s}. \tag{2.94}
\end{aligned}$$

Evaluating the first few operators,

$$\hat{C}_{0,1} = -\frac{\hat{H}_0}{F_0}, \tag{2.95}$$

$$\hat{C}_{0,2} = \frac{1}{2F_0} \left( \frac{F_1 + G_0}{F_0} \hat{H}_0 - \hat{H}_1 \right) + \frac{1}{\mathcal{L}_0 h_{X0} F_0} (\mathcal{L}_1 h_{X0} + \mathcal{L}_0 h_{X1}) \hat{H}_0, \tag{2.96}$$

$$\hat{C}_{1,1} = -\frac{G_0}{F_0} - \frac{\mathcal{L}_1 h_{X0} + \mathcal{L}_0 h_{X1}}{\mathcal{L}_0 h_{X0}}, \tag{2.97}$$

$$\begin{aligned}
\hat{C}_{1,2} = & \frac{1}{2F_0} \left( \frac{F_1 + G_0}{F_0} G_0 - G_1 - \hat{H}_0 \right) \\
& - \frac{1}{\mathcal{L}_0 h_{X0}} \left( \begin{array}{c} \mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1} + \mathcal{L}_0 h_{X2}, \\ -(\mathcal{L}_1 h_{X0} + \mathcal{L}_0 h_{X1}) (G_0/F_0 + (\mathcal{L}_1 h_{X0} + \mathcal{L}_0 h_{X1})/\mathcal{L}_0 h_{X0}) \end{array} \right). \tag{2.98}
\end{aligned}$$

From this,  $\xi_X$  for the non-resonant case is written as,

$$\xi_X = \frac{B_0}{\mathcal{L}_0 h_{X0}} \left( \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} \right) \alpha_0(Y) + \frac{B_0}{\mathcal{L}_0 h_{X0}} \left( \sum_{n=0}^{\infty} X^n \hat{C}_{1,n} \right) \alpha_1(Y), \tag{2.99}$$

$$\begin{aligned}
\xi_X = & \frac{B_0}{\mathcal{L}_0 h_{X0}} \left[ X \left( -\frac{\hat{H}_0}{F_0} \right) + \dots \right] \alpha_0(Y) \\
& + \frac{B_0}{\mathcal{L}_0 h_{X0}} \left[ 1 + X \left( -\frac{G_0}{F_0} - \frac{\mathcal{L}_1 h_{X0} + \mathcal{L}_0 h_{X1}}{\mathcal{L}_0 h_{X0}} \right) + \dots \right] \alpha_1(Y). \tag{2.100}
\end{aligned}$$

#### 2.5.4.8 $\xi_Y$

Equation (2.39) states

$$\mathcal{L} h_Y \xi_Y = B_0 \frac{\partial b_z}{\partial Y}.$$

We solve for  $\xi_Y$  in the same manner as for  $\xi_X$ , led by the form of  $b_z$  and equation (2.39) to expand  $\xi_Y$  as

$$\xi_Y = X^\sigma \frac{B_0}{\mathcal{L}_0 h_{Y0}} \left( \sum_{n=0}^{\infty} X^n \hat{J}_{0,n} \right) \alpha'_0(Y) + X^\sigma \frac{B_0}{\mathcal{L}_0 h_{Y0}} \left( \sum_{n=0}^{\infty} X^n \hat{J}_{1,n} \right) \alpha'_1(Y), \quad (2.101)$$

where  $\alpha'_i(Y) = d\alpha_i(Y)/dY$  and  $\sigma$  is a new parameter to be determined. Once again,  $b_z$  does not involve any logarithms so we do not assume them in our expansion of  $\xi_Y$ .

Substituting in equation (2.39),

$$\begin{aligned} & \left( \sum_{n=0}^{\infty} X^n \mathcal{L}_n \right) \left( \sum_{n=0}^{\infty} X^n h_{Yn} \right) X^\sigma \frac{B_0}{\mathcal{L}_0 h_{Y0}} \times \\ & \quad \left( \left( \sum_{n=0}^{\infty} X^n \hat{J}_{0,n} \right) \alpha'_0(Y) + \left( \sum_{n=0}^{\infty} X^n \hat{J}_{1,n} \right) \alpha'_1(Y) \right) \\ & = B_0 \left( \sum_{n=0}^{\infty} X^n \hat{A}_{0,n} \right) \alpha'_0(Y) + B_0 \left( \sum_{n=0}^{\infty} X^n \hat{A}_{1,n} \right) \alpha'_1(Y). \end{aligned} \quad (2.102)$$

The lowest order, non-zero term on the RHS of the above equation is

$$B_0 \alpha'_0(Y) X^0,$$

which must match with the lowest order, non-zero term on the LHS. Requiring that one of  $\hat{J}_{0,0}$  and  $\hat{J}_{1,0}$  be non-zero, the lowest order term on the LHS is

$$B_0 \left( \hat{J}_{0,0} \alpha'_0 + \hat{J}_{1,0} \alpha'_1 \right) X^\sigma,$$

so matching the LHS and RHS gives  $\sigma = 0$  and

$$\hat{J}_{0,0} = 1, \quad (2.103)$$

$$\hat{J}_{1,0} = 0. \quad (2.104)$$

Remaining operators are now obtained by matching terms in  $X^r \alpha'_i$ . This gives,

$$\begin{aligned} & \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Yt} \hat{J}_{i,s} = \mathcal{L}_0 h_{Y0} \hat{A}_{i,r} \\ \Rightarrow & \mathcal{L}_0 h_{Y0} \hat{J}_{i,r} + \sum_{s=0}^{r-1} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Yt} \hat{J}_{i,s} = \mathcal{L}_0 h_{Y0} \hat{A}_{i,r} \end{aligned}$$

$$\Rightarrow \hat{J}_{i,r} = \hat{A}_{i,r} - \frac{1}{\mathcal{L}_0 h_{Y0}} \sum_{s=0}^{r-1} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Yt} \hat{J}_{i,s}. \quad (2.105)$$

Evaluating the first few operators,

$$\hat{J}_{0,1} = -\frac{\mathcal{L}_1 h_{Y0} + \mathcal{L}_0 h_{Y1}}{\mathcal{L}_0 h_{Y0}}, \quad (2.106)$$

$$\hat{J}_{0,2} = -\frac{\hat{H}_0}{2F_0} - \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1} + \mathcal{L}_0 h_{Y2}}{\mathcal{L}_0 h_{Y0}} + \left( \frac{\mathcal{L}_1 h_{Y0} + \mathcal{L}_0 h_{Y1}}{\mathcal{L}_0 h_{Y0}} \right)^2, \quad (2.107)$$

$$\hat{J}_{1,1} = 1, \quad (2.108)$$

$$\hat{J}_{1,2} = -\frac{G_0}{2F_0} - \frac{\mathcal{L}_1 h_{Y0} + \mathcal{L}_0 h_{Y1}}{\mathcal{L}_0 h_{Y0}}. \quad (2.109)$$

From this,  $\xi_Y$  for the non-resonant case is written as

$$\xi_Y = \frac{B_0}{\mathcal{L}_0 h_{Y0}} \left( \left( \sum_{n=0}^{\infty} X^n \hat{J}_{0,n} \right) \alpha'_0(Y) + \left( \sum_{n=0}^{\infty} X^n \hat{J}_{1,n} \right) \alpha'_1(Y) \right), \quad (2.110)$$

$$\begin{aligned} \xi_Y &= \frac{B_0}{\mathcal{L}_0 h_{Y0}} \left[ 1 + X \left( -\frac{\mathcal{L}_1 h_{Y0} + \mathcal{L}_0 h_{Y1}}{\mathcal{L}_0 h_{Y0}} \right) + \dots \right] \alpha'_0(Y) \\ &\quad + \frac{B_0}{\mathcal{L}_0 h_{Y0}} \left[ X + X^2 \left( -\frac{G_0}{2F_0} - \frac{\mathcal{L}_1 h_{Y0} + \mathcal{L}_0 h_{Y1}}{\mathcal{L}_0 h_{Y0}} \right) + \dots \right] \alpha'_1(Y). \end{aligned} \quad (2.111)$$

### 2.5.5 Resonant Solution

In the resonant case, there is at least one surface at which the frequency of the system,  $\omega$ , matches the natural field-line Alfvén frequency,  $\omega_A = k_z v_A = k_z B_0 / \sqrt{\mu_0 \rho_e}$ . In order to investigate the solution in the vicinity of this surface, the origin of  $X$  is chosen so that  $\omega_A(X=0) = \omega$ , i.e. so that the resonant surface is positioned at  $X=0$ .

Referring back to (2.52), resonance means that  $\mathcal{L}_0 = \mathcal{L}(0) = 0$ , so  $F_0 = 0$  by (2.59), and  $\hat{H}_0 = 0$  by (2.63). Therefore, equations (2.67) and (2.68) are automatically satisfied. We have, however, required that  $\partial \rho_e / \partial X \neq 0$  at  $X=0$ , so,  $\mathcal{L}_1 \neq 0$  by (2.53), giving  $F_1 \neq 0$  by (2.60). Also note that, for the resonant case,

$$F_1 = h_{Y0}^3 h_{X0} \mathcal{L}_1 = -G_0, \quad (2.112)$$

which will be useful for simplifying in equations (2.69) to (2.72).

**2.5.5.1 Value of  $\beta_0$** 

If  $\sigma \notin \{0, 2\}$  then (2.69) gives  $\beta_0 = 0$ . If  $\sigma \in \{0, 2\}$  then (2.70) gives  $\beta_0 = 0$ . Therefore,  $\beta_0 = 0 \forall \sigma$ .

**2.5.5.2 Choosing  $\sigma$  for General Solution**

Putting  $\beta_0 = 0$  into (2.70) gives  $\sigma(\sigma - 2)\alpha_0 = 0$ . We require  $\alpha_0 \neq 0$  for a non-trivial solution, which implies  $\sigma \in \{0, 2\}$ . Since the possible values for  $\sigma$  are separated by an integer, the lowest value,  $\sigma = 0$ , corresponds to the 2-parameter general solution.

**2.5.5.3 Value of  $\beta_1$** 

If  $\sigma \in \{0, 2\}$ , as required by (2.5.5.2), then evaluating (2.71) for  $n = 0$  implies  $\beta_1 = 0$ .

**2.5.5.4 Value of  $\alpha_1$** 

Evaluating (2.72) for  $n = 0$  with  $\sigma = \beta_0 = 0$  gives  $\alpha_1 = 0$  in the general solution.

**2.5.5.5 Formula for  $\beta_2$** 

Evaluating (2.72) for  $n = 1$  with  $\sigma = \beta_0 = \beta_1 = \alpha_1 = 0$  gives

$$\beta_2 = -\frac{\hat{H}_1}{2F_1}\alpha_0. \quad (2.113)$$

**2.5.5.6 Recurrence Relation for  $\beta_m$ ,  $m \geq 3$** 

Putting  $\sigma = 0$  into equation (2.71) gives, for  $n \geq 2$ ,

$$\begin{aligned} & F_1(n-1)(n+1)\beta_{n+1} + \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \beta_s = 0 \\ \Rightarrow & \beta_{n+1} = -\frac{1}{(n-1)(n+1)F_1} \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \beta_s \\ \Rightarrow & \beta_m = -\frac{1}{m(m-2)F_1} \sum_{s=0}^{m-1} \left( s(s-1)F_{m-s+1} + sG_{m-s} + \hat{H}_{m-s-1} \right) \beta_s. \end{aligned} \quad (2.114)$$

Thus, all  $\beta_m$ ,  $m \geq 3$ , are determined in terms of  $\beta_2$ , which is itself determined from  $\alpha_0$ .

### 2.5.5.7 Recurrence Relation for $\alpha_m$ , $m \geq 3$

Putting  $\sigma = 0$  into equation (2.72) gives, for  $n \geq 2$ ,

$$(n+1)(n-1)F_1\alpha_{n+1} + \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \alpha_s \\ + 2nF_1\beta_{n+1} + \sum_{s=0}^n \left( (2s-1)F_{n-s+2} + G_{n-s+1} \right) \beta_s = 0$$

$$\Rightarrow \alpha_{n+1} = -\frac{1}{(n-1)(n+1)F_1} \left[ \begin{array}{c} 2nF_1\beta_{n+1} \\ + \sum_{s=0}^n \left( s(s-1)F_{n-s+2} + sG_{n-s+1} + \hat{H}_{n-s} \right) \alpha_s \\ + \sum_{s=0}^n \left( (2s-1)F_{n-s+2} + G_{n-s+1} \right) \beta_s \end{array} \right] \\ \Rightarrow \alpha_m = -\frac{1}{m(m-2)F_1} \left[ \begin{array}{c} 2(m-1)F_1\beta_m \\ + \sum_{s=0}^{m-1} \left( s(s-1)F_{m-s+1} + sG_{m-s} + \hat{H}_{m-s-1} \right) \alpha_s \\ + \sum_{s=0}^{m-1} \left( (2s-1)F_{m-s+1} + G_{m-s} \right) \beta_s \end{array} \right]. \quad (2.115)$$

This determines all  $\alpha_m$ ,  $m \geq 3$  in terms of  $\alpha_0$  and  $\alpha_2$ . Since all non-zero  $\beta_j$  are determined in terms of  $\alpha_0$ , the general 2-parameter solution has been obtained.

### 2.5.5.8 General Solution as Sum of 1-Parameter Solutions

As in the non-resonant case, the general solution may be written as a sum of two independent 1-parameter solutions. The first of these is obtained by considering  $\alpha_0(Y) \neq 0$  with  $\alpha_2(Y) = 0$ . Since all non-zero  $\beta_j$  are determined in terms of  $\alpha_0$ , this solution will contain all the terms involving logarithms.

Writing

$$\beta_m = \hat{B}_m\beta_2, \quad (2.116)$$

we have

$$\hat{B}_0 = 0, \quad (2.117)$$

$$\hat{B}_1 = 0, \quad (2.118)$$

$$\hat{B}_2 = 1. \quad (2.119)$$

Substituting equation (2.116) into (2.114) gives a recurrence relation for the higher operators,

$$\hat{B}_m = -\frac{1}{m(m-2)F_1} \sum_{s=0}^{m-1} \left( s(s-1)F_{m-s+1} + sG_{m-s} + \hat{H}_{m-s-1} \right) \hat{B}_s. \quad (2.120)$$

Using this to evaluate the first few operators,

$$\hat{B}_3 = -\frac{2}{3F_1} (F_2 + G_1), \quad (2.121)$$

$$\hat{B}_4 = -\frac{1}{8F_1} \left( 2F_3 + 2G_2 + \hat{H}_1 - \frac{2}{F_1} (2F_2 + G_1) (F_2 + G_1) \right). \quad (2.122)$$

The remaining part of the solution is obtained by writing

$$\alpha_m = \hat{A}_{0,m} \alpha_0, \quad (2.123)$$

with

$$\hat{A}_{0,0} = 1, \quad (2.124)$$

$$\hat{A}_{0,1} = 0, \quad (2.125)$$

$$\hat{A}_{0,2} = 0, \quad (2.126)$$

and evaluating higher  $\hat{A}_{0,m}$  by recurrence relation.

The second independent 1-parameter solution is obtained by taking  $\alpha_0(Y) = 0$  with  $\alpha_2(Y) \neq 0$ . As has already been noted, all  $\beta_m$  are zero in this solution. This means that the complete solution is obtained by writing,

$$\alpha_m = \hat{A}_{2,m} \alpha_2, \quad (2.127)$$

with,

$$\hat{A}_{2,0} = 0, \quad (2.128)$$

$$\hat{A}_{2,1} = 0, \quad (2.129)$$

$$\hat{A}_{2,2} = 1, \quad (2.130)$$

and evaluating higher  $\hat{A}_{2,m}$  by recurrence relation.

The recurrence relation for higher  $\hat{A}_{i,m}$  is obtained by substituting

$$\alpha_m = \hat{A}_{i,m} \alpha_i, \quad (2.131)$$

and

$$\beta_m = -\delta_{0,i} \frac{\hat{B}_m \hat{H}_1}{2F_1} \alpha_i, \quad (2.132)$$

into equation (2.115), where

$$\delta_{0,0} = 1, \quad (2.133)$$

$$\delta_{0,i} = 0, \quad i \neq 0. \quad (2.134)$$

The resulting recurrence relation is

$$\hat{A}_{i,m} = -\frac{1}{m(m-2)F_1} \left[ \begin{array}{c} -\delta_{0,i}(m-1)\hat{B}_m\hat{H}_1 \\ + \sum_{s=0}^{m-1} (s(s-1)F_{m-s+1} + sG_{m-s} + \hat{H}_{m-s-1}) \hat{A}_{i,s} \\ - \delta_{0,i} \sum_{s=0}^{m-1} ((2s-1)F_{m-s+1} + G_{m-s}) \hat{B}_s \hat{H}_1 / 2F_1 \end{array} \right], \quad (2.135)$$

valid for  $m \geq 3$ . Note that the recurrence relation and generating conditions for  $\hat{A}_{2,m}$  are identical to those for  $\hat{B}_m$ . Therefore,

$$\hat{A}_{2,m} = \hat{B}_m. \quad (2.136)$$

Evaluating the first few operators,

$$\hat{A}_{0,3} = \frac{1}{3F_1} \left( \frac{1}{6F_1} (F_2 - 5G_1) \hat{H}_1 - \hat{H}_2 \right), \quad (2.137)$$

$$\hat{A}_{2,3} = \hat{B}_3 = -\frac{2}{3F_1} (F_2 + G_1), \quad (2.138)$$

$$\hat{A}_{2,4} = \hat{B}_4 = -\frac{1}{8F_1} \left( 2F_3 + 2G_2 + \hat{H}_1 - \frac{2}{F_1} (2F_2 + G_1) (F_2 + G_1) \right). \quad (2.139)$$

Finally, the solution to the resonant problem is written as,

$$b_z = \sum_{n=0}^{\infty} X^n \alpha_n + \ln(X) \sum_{n=0}^{\infty} X^n \beta_n, \quad (2.140)$$

$$\begin{aligned} b_z &= \left( \sum_{n=0}^{\infty} X^n \hat{A}_{0,n} - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{B}_n \hat{H}_1 \right) \alpha_0(Y) \\ &+ \left( \sum_{n=0}^{\infty} X^n \hat{A}_{2,n} \right) \alpha_2(Y), \end{aligned} \quad (2.141)$$

$$\begin{aligned} b_z &= \left[ \begin{array}{l} 1 + X^3 \left( (F_2 - 5G_1) \hat{H}_1 / 6F_1 - \hat{H}_2 \right) / 3F_1 + \dots \\ - \ln(X) \left[ X^2 - X^3 2(F_2 + G_1) / 3F_1 + \dots \right] \hat{H}_1 / 2F_1 \end{array} \right] \alpha_0(Y) \\ &+ \left[ X^2 - X^3 2(F_2 + G_1) / 3F_1 + \dots \right] \alpha_2(Y). \end{aligned} \quad (2.142)$$

2.5.5.9  $\xi_X$ 

Equation (2.38) states

$$\mathcal{L}h_X \xi_X = B_0 \frac{\partial b_z}{\partial X}.$$

In the resonant case,

$$\begin{aligned} \frac{\partial b_z}{\partial X} &= \left( \sum_{n=0}^{\infty} X^{n-1} \left( n\hat{A}_{0,n} - \frac{1}{2F_1} \hat{B}_n \hat{H}_1 \right) - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} nX^{n-1} \hat{B}_n \hat{H}_1 \right) \alpha_0(Y) \\ &\quad + \left( \sum_{n=0}^{\infty} nX^{n-1} \hat{A}_{2,n} \right) \alpha_2(Y). \end{aligned} \quad (2.143)$$

Led by the form of  $\partial b_z / \partial X$  and equation (2.38), we let  $\xi_X$  take the form

$$\begin{aligned} \xi_X &= X^\sigma \frac{B_0}{\mathcal{L}_1 h_{X0}} \left( -\frac{1}{2F_1} \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} - \frac{1}{F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{D}_n \hat{H}_1 \right) \alpha_0(Y) \\ &\quad + X^\sigma \frac{B_0}{\mathcal{L}_1 h_{X0}} \left( 2 \sum_{n=0}^{\infty} X^n \hat{C}_{2,n} \right) \alpha_2(Y), \end{aligned} \quad (2.144)$$

where  $\sigma$  is once again a free parameter.

Substituting in equation (2.38),

$$\begin{aligned} &\left( \sum_{n=0}^{\infty} X^n \mathcal{L}_n \right) \left( \sum_{n=0}^{\infty} X^n h_{Xn} \right) X^\sigma \frac{B_0}{\mathcal{L}_1 h_{X0}} \times \\ &\quad \left[ \left( -\frac{1}{2F_1} \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} - \frac{1}{F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{D}_n \hat{H}_1 \right) \alpha_0(Y) \right. \\ &\quad \left. + \left( 2 \sum_{n=0}^{\infty} X^n \hat{C}_{2,n} \right) \alpha_2(Y) \right] \\ &= B_0 \left( \sum_{n=0}^{\infty} X^{n-1} \left( n\hat{A}_{0,n} - \frac{1}{2F_1} \hat{B}_n \hat{H}_1 \right) - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} nX^{n-1} \hat{B}_n \hat{H}_1 \right) \alpha_0(Y) \\ &\quad + B_0 \left( \sum_{n=0}^{\infty} nX^{n-1} \hat{A}_{2,n} \right) \alpha_2(Y). \end{aligned} \quad (2.145)$$

The lowest order, non-zero term on the RHS is

$$B_0 \left[ \left( -\frac{1}{2F_1} \hat{H}_1 - \frac{1}{F_1} \ln(X) \hat{H}_1 \right) \alpha_0(Y) + 2\alpha_2(Y) \right] X^1,$$

which must match the lowest order, non-zero term on the LHS. Requiring that one of  $\hat{C}_{0,0}$ ,  $\hat{C}_{2,0}$

and  $\hat{D}_0$  be non-zero, the lowest order term on the LHS is

$$B_0 \left[ \left( -\frac{1}{2F_1} \hat{C}_{0,0} - \frac{1}{F_1} \ln(X) \hat{D}_0 \hat{H}_1 \right) \alpha_0(Y) + 2\hat{C}_{2,0} \alpha_2(Y) \right] X^{\sigma+1},$$

so matching the LHS and RHS gives  $\sigma = 0$  and

$$\hat{C}_{0,0} = \hat{H}_1, \quad (2.146)$$

$$\hat{C}_{2,0} = 1, \quad (2.147)$$

$$\hat{D}_0 = 1. \quad (2.148)$$

Remaining operators are now obtained by matching terms. Matching terms in  $X^r \alpha_0(Y)$  gives

$$\begin{aligned} & -\frac{1}{2\mathcal{L}_1 h_{X0} F_1} \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{0,s} = (r+1) \hat{A}_{0,r+1} - \frac{1}{2F_1} \hat{B}_{r+1} \hat{H}_1 \\ \Rightarrow \hat{C}_{0,r-1} &= \hat{B}_{r+1} \hat{H}_1 - 2F_1(r+1) \hat{A}_{0,r+1} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{r-2} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{0,s} \\ \Rightarrow \hat{C}_{0,m} &= \hat{B}_{m+2} \hat{H}_1 - 2F_1(m+2) \hat{A}_{0,m+2} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{m-1} \sum_{t=0}^{m+1-s} \mathcal{L}_{m+1-s-t} h_{Xt} \hat{C}_{0,s}. \end{aligned} \quad (2.149)$$

Matching terms in  $X^r \alpha_2$  gives

$$\begin{aligned} & \frac{2}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{2,s} = (r+1) \hat{A}_{2,r+1} \\ \Rightarrow \hat{C}_{2,r-1} &= \frac{r+1}{2} \hat{A}_{2,r+1} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{r-2} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{C}_{2,s} \\ \Rightarrow \hat{C}_{2,m} &= \frac{m+2}{2} \hat{A}_{2,m+2} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{m-1} \sum_{t=0}^{m+1-s} \mathcal{L}_{m+1-s-t} h_{Xt} \hat{C}_{2,s}. \end{aligned} \quad (2.150)$$

Matching terms in  $\ln(X) X^r \hat{H}_1 \alpha_0$  gives

$$\begin{aligned} & \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{D}_s = \frac{r+1}{2} \hat{B}_{r+1} \\ \Rightarrow \hat{D}_{r-1} &= \frac{r+1}{2} \hat{B}_{r+1} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{r-2} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Xt} \hat{D}_s \\ \Rightarrow \hat{D}_m &= \frac{m+2}{2} \hat{B}_{m+2} - \frac{1}{\mathcal{L}_1 h_{X0}} \sum_{s=0}^{m-1} \sum_{t=0}^{m+1-s} \mathcal{L}_{m+1-s-t} h_{Xt} \hat{D}_s. \end{aligned} \quad (2.151)$$

The recurrence relations for  $\hat{C}_{2,m}$  and  $\hat{D}_m$  are identical. Since  $\hat{A}_{2,m} = \hat{B}_m$  and  $\hat{C}_{2,0} = \hat{D}_0$  this gives

$$\hat{C}_{2,m} = \hat{D}_m. \quad (2.152)$$

Evaluating the first few operators,

$$\hat{C}_{0,1} = \frac{1}{F_1} (-F_2 + G_1) \hat{H}_1 + 2\hat{H}_2 - \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \hat{H}_1, \quad (2.153)$$

$$\hat{C}_{2,1} = \hat{D}_1 = -\frac{1}{F_1} (F_2 + G_1) - \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}}, \quad (2.154)$$

$$\begin{aligned} \hat{C}_{2,2} = \hat{D}_2 = & -\frac{1}{4F_1} \left( 2F_3 + 2G_2 + \hat{H}_1 - \frac{2}{F_1} (2F_2 + G_1) (F_2 + G_1) \right) \\ & - \frac{\mathcal{L}_3 h_{X0} + \mathcal{L}_2 h_{X1} + \mathcal{L}_1 h_{X2}}{\mathcal{L}_1 h_{X0}} \\ & + \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \left( \frac{1}{F_1} (F_1 + G_1) + \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \right), \end{aligned} \quad (2.155)$$

from which,  $\xi_X$  for the resonant case is

$$\begin{aligned} \xi_X = & \frac{B_0}{\mathcal{L}_1 h_{X0}} \left( -\frac{1}{2F_1} \sum_{n=0}^{\infty} X^n \hat{C}_{0,n} - \frac{1}{F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{D}_n \hat{H}_1 \right) \alpha_0(Y) \\ & + \frac{B_0}{\mathcal{L}_1 h_{X0}} \left( 2 \sum_{n=0}^{\infty} X^n \hat{C}_{2,n} \right) \alpha_2(Y), \end{aligned} \quad (2.156)$$

$$\begin{aligned} \xi_X = & \frac{B_0}{\mathcal{L}_1 h_{X0}} \left[ -\frac{1}{2F_1} \left[ \hat{H}_1 + X \left( \frac{(G_1 - F_2)}{F_1} \hat{H}_1 + 2\hat{H}_2 - \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \hat{H}_1 \right) + \dots \right] \right. \\ & \left. - \frac{1}{F_1} \ln(X) \left[ 1 + X \left( -\frac{(F_2 + G_1)}{F_1} - \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \right) + \dots \right] \hat{H}_1 \right] \alpha_0(Y) \\ & + \frac{2B_0}{\mathcal{L}_1 h_{X0}} \left[ 1 + X \left( -\frac{1}{F_1} (F_2 + G_1) - \frac{\mathcal{L}_2 h_{X0} + \mathcal{L}_1 h_{X1}}{\mathcal{L}_1 h_{X0}} \right) + \dots \right] \alpha_2(Y). \end{aligned} \quad (2.157)$$

### 2.5.5.10 $\xi_Y$

Equation (2.39) states

$$\mathcal{L} h_Y \xi_Y = B_0 \frac{\partial b_z}{\partial Y}.$$

We solve for  $\xi_Y$  in the same manner as for  $\xi_X$ , led by the form of  $b_z$  and equation (2.39) to expand  $\xi_Y$  as

$$\xi_Y = X^\sigma \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left( \sum_{n=0}^{\infty} X^n \hat{J}_{0,n} - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{K}_n \hat{H}_1 \right) \alpha'_0(Y)$$

$$+X^\sigma \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left( \sum_{n=0}^{\infty} X^n \hat{J}_{2,n} \right) \alpha'_2(Y), \quad (2.158)$$

where  $\alpha'_i(Y) = d\alpha_i(Y)/dY$  and  $\sigma$  is a new parameter to be determined.

Substituting in equation (2.39),

$$\begin{aligned} & \left( \sum_{n=0}^{\infty} X^n \mathcal{L}_n \right) \left( \sum_{n=0}^{\infty} X^n h_{Yn} \right) X^\sigma \frac{B_0}{\mathcal{L}_1 h_{Y0}} \\ & \times \left[ \left( \sum_{n=0}^{\infty} X^n \hat{J}_{0,n} - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{K}_n \hat{H}_1 \right) \alpha'_0(Y) + \left( \sum_{n=0}^{\infty} X^n \hat{J}_{2,n} \right) \alpha'_2(Y) \right] \\ & = B_0 \left( \sum_{n=0}^{\infty} X^n \hat{A}_{0,n} - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} X^n \hat{B}_n \hat{H}_1 \right) \alpha'_0(Y) \\ & + B_0 \left( \sum_{n=0}^{\infty} X^n \hat{A}_{2,n} \right) \alpha'_2(Y). \end{aligned} \quad (2.159)$$

The lowest order, non-zero term on the RHS of the above equation is

$$B_0 \alpha'_0(Y) X^0,$$

which must match with the lowest order, non-zero term on the LHS. Requiring that one of  $\hat{J}_{0,0}$ ,  $\hat{J}_{1,0}$  and  $\hat{K}_n$  be non-zero, the lowest order term on the LHS is

$$B_0 \left[ \left( \hat{J}_{0,0} - \frac{1}{2F_1} \ln(X) \hat{K}_0 \hat{H}_1 \right) \alpha'_0(Y) + \hat{J}_{2,0} \alpha'_2(Y) \right] X^{\sigma+1},$$

so matching the LHS and RHS gives  $\sigma = -1$  and

$$\hat{J}_{0,0} = 1, \quad (2.160)$$

$$\hat{J}_{2,0} = 0, \quad (2.161)$$

$$\hat{K}_0 = 0. \quad (2.162)$$

Remaining operators are now obtained by matching terms. Matching terms in  $X^{r-1} \alpha'_i(Y)$  gives

$$\begin{aligned} & \frac{1}{\mathcal{L}_1 h_{Y0}} \sum_{s=0}^r \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Yt} \hat{J}_{i,s} = \hat{A}_{i,r-1} \\ \Rightarrow & \hat{J}_{i,r-1} = \hat{A}_{i,r-1} - \frac{1}{\mathcal{L}_1 h_{Y0}} \sum_{s=0}^{r-2} \sum_{t=0}^{r-s} \mathcal{L}_{r-s-t} h_{Yt} \hat{J}_{i,s} \end{aligned}$$

$$\Rightarrow \hat{J}_{i,m} = \hat{A}_{i,m} - \frac{1}{\mathcal{L}_1 h_{Y0}} \sum_{s=0}^{m-1} \sum_{t=0}^{m+1-s} \mathcal{L}_{m+1-s-t} h_{Yt} \hat{J}_{i,s}. \quad (2.163)$$

Matching terms in  $X^{r-1} \ln(X) \alpha'_0(Y)$  produces an equation of the same form as (2.163), in which  $\hat{K}_m$  replaces  $\hat{J}_{i,m}$  and  $\hat{B}_m$  replaces  $\hat{A}_{i,m}$ . Since  $\hat{B}_m = \hat{A}_{2,m}$  this gives

$$\hat{K}_m = \hat{J}_{2,m}. \quad (2.164)$$

Evaluating the first few operators,

$$\hat{J}_{0,1} = -\frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}}, \quad (2.165)$$

$$\hat{J}_{0,2} = -\frac{\mathcal{L}_3 h_{Y0} + \mathcal{L}_2 h_{Y1} + \mathcal{L}_1 h_{Y2}}{\mathcal{L}_1 h_{Y0}} + \left( \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}} \right)^2, \quad (2.166)$$

$$\hat{J}_{2,1} = \hat{K}_1 = 0, \quad (2.167)$$

$$\hat{J}_{2,2} = \hat{K}_2 = 1, \quad (2.168)$$

$$\hat{J}_{2,3} = \hat{K}_3 = -\frac{2}{3F_1} (F_2 + G_1) - \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}}. \quad (2.169)$$

From this,  $\xi_Y$  for the resonant case is

$$\begin{aligned} \xi_Y &= \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left( \sum_{n=0}^{\infty} X^{n-1} \hat{J}_{0,n} - \frac{1}{2F_1} \ln(X) \sum_{n=0}^{\infty} X^{n-1} \hat{K}_n \hat{H}_1 \right) \alpha'_0(Y) \\ &+ \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left( \sum_{n=0}^{\infty} X^{n-1} \hat{J}_{2,n} \right) \alpha'_2(Y), \end{aligned} \quad (2.170)$$

$$\begin{aligned} \xi_Y &= \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left[ \begin{aligned} &\left( \frac{1}{X} - \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}} + \dots \right) \\ &- \frac{1}{2F_1} \ln(X) \left( X + X^2 \left( -\frac{2(F_2 + G_1)}{3F_1} - \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}} \right) + \dots \right) \hat{H}_1 \end{aligned} \right] \alpha'_0(Y) \\ &+ \frac{B_0}{\mathcal{L}_1 h_{Y0}} \left( X + X^2 \left( -\frac{2(F_2 + G_1)}{3F_1} - \frac{\mathcal{L}_2 h_{Y0} + \mathcal{L}_1 h_{Y1}}{\mathcal{L}_1 h_{Y0}} \right) + \dots \right) \alpha'_2(Y). \end{aligned} \quad (2.171)$$

## 2.5.6 Summary of Leading Order Behaviours

### 2.5.6.1 Non-Resonant Case

In the non-resonant case,

$$\begin{aligned} b_z &= \alpha_0(Y) + \alpha_1(Y)X + O(X^2), \\ \xi_X &= \frac{B_0}{\mathcal{L}_0 h_{X0}} \alpha_1 + O(X), \end{aligned}$$

$$\xi_Y = \frac{B_0}{\mathcal{L}_0 h_{Y0}} \alpha'_0 + O(X).$$

### 2.5.6.2 Resonant Case

In the resonant case,

$$\begin{aligned} b_z &= \alpha_0(Y) + \alpha_2(Y)X^2 + O(X^3) + \beta_2(Y) \ln(X)X^2 + \ln(X)O(X^3), \\ \xi_X &= \frac{B_0}{\mathcal{L}_1 h_{X0}} (2\alpha_2 + \beta_2)X^0 + O(X^1) + \frac{2B_0}{\mathcal{L}_1 h_{X0}} \beta_2 \ln(X)X^0 + \ln(X)O(X^1), \\ \xi_Y &= \frac{B_0}{\mathcal{L}_1 h_{Y0}} \alpha'_0 X^{-1} + O(X^0) + \frac{B_0}{\mathcal{L}_1 h_{Y0}} \beta'_2 \ln(X)X^1 + \ln(X)O(X^2), \end{aligned}$$

where

$$\beta_2(Y) = -\frac{h_{X0}}{2h_{Y0}^3} \left( (h_{Y0}h'_{X0} - h_{X0}h'_{Y0}) \frac{d}{dY} + h_{X0}h_{Y0} \frac{d^2}{dY^2} \right) \alpha_0(Y).$$

### 2.5.6.3 Solution on Resonant Surface

The solution on the resonant surface can be examined by taking the limit  $X \rightarrow 0$ , which gives

$$b_z = \alpha_0(Y), \tag{2.172}$$

$$\xi_X = \frac{2B_0}{\mathcal{L}_1 h_{X0}} \beta_2(Y) \ln(X), \tag{2.173}$$

$$\xi_Y = \frac{B_0}{\mathcal{L}_1 h_{Y0}} \alpha'_0(Y) X^{-1}. \tag{2.174}$$

Thus,  $\alpha_0(Y)$  represents  $b_z$  evaluated on the resonant contour. This is proportional to the magnetic-pressure perturbation and is associated with the fast wave. The nature of the singularity is such that the dominant displacement is  $\xi_Y$ , leading to a velocity perturbation

$$u_Y = -i \left( \frac{B_0}{\mu_0 \omega [\partial \rho / \partial X]_{X=0}} \right) \left( \frac{1}{h_{Y0}} \left[ \frac{\partial b_z}{\partial Y} \right]_{X=0} \right) X^{-1}. \tag{2.175}$$

This is the velocity perturbation of the resonant Alfvén wave, which is the dominant feature for late times. Note the dependence on  $db_z/(h_{Y0}dY)$ , which is equivalent to  $\tilde{d}b_z/\tilde{d}s$  in the numerical simulations. This confirms that the spatial form of the fast wave leads to amplitude variations of the resonant Alfvén wave. We also note a dependence on the frequency,  $\omega$ , and the density gradient across the resonance,  $[\partial \rho / \partial X]_{X=0}$ .

## 2.6 Discussion

We have established a mathematical basis by which field-line resonance may be understood when field-line Alfvén eigenfrequencies vary in two dimensions, and we have presented computer simulations showing the main features of this fundamental MHD process. Together, these results confirm that much of our intuitive understanding applies to such 2D equilibria: particularly, energy is deposited as a phase-mixing Alfvén wave where the field-line eigenfrequency matches the driving frequency. This work fills a gap that has previously existed between mathematical theory (e.g. Southwood (1974)), simulations (including the impulsively-driven 2D simulations of Lee et al. (2000) and Terradas et al. (2008)) and observations of field-line resonance in non-axisymmetric systems such as magnetospheric flanks and the magnetotail. We have also shown how the 1D phase-mixing length (Mann et al., 1995) can be generalised for higher dimensions.

An important feature of 2D field-line resonance is the concept of ‘imprinting’. Here, a resonant Alfvén wave exhibits amplitude variations along contours of Alfvén frequency, and these can be used to reveal the spatial form of the fast wave, because the Alfvén wave velocity perturbation correlates strongly with the magnetic pressure gradient of the fast wave. Intuitively, one can think of a resonant field-line receiving a ‘push’ from the fast wave during every period. These pushes are large where the magnetic pressure gradient is large, and small where the magnetic pressure gradient is small. After several cycles, those field-lines which have received the largest pushes have the largest Alfvén oscillations. Alternatively, one can consider the variation of fast wave energy at the resonant surface. This is reflected in the energy available through resonance, leading to amplitude variations of the Alfvén wave. This is in contrast to decoupled Alfvén waves, which are invariant along contours of Alfvén frequency.

In the magnetosphere, the correlation between the amplitude of the Alfvén wave and the amplitude of the magnetic pressure force offers a means of probing the magnetosphere through magnetoseismology. Observations of ULF magnetic pulsations over different magnetic latitudes and local times provide a spatial picture of the Alfvén wave, which, in turn, reveals the structure of the magnetospheric fast wave. This connection has already been exploited in weaker forms, for example, the low power of ULF waves at local noon and the dominance of antisunward azimuthal phase speeds led the community to consider antisunward propagating fast waves as the dominant driver (Anderson et al., 1990; Samson et al., 1992). More quantitatively, Wright and Rickard (1995) showed that a displacement pulse running along the magnetopause excites resonant Alfvén waves with an azimuthal phase velocity strictly equal to that of the boundary pulse. The present study suggests, however, that the correlation can be exploited more generally, also using features in amplitude.

This work is readily applied to a coronal loop with a continuous profile of Alfvén speed. For a monolithic, cylindrical loop (with a uniform magnetic field aligned along the axis of the loop

and density decreasing radially from the centre of the loop), the total pressure perturbation of the kink wave has an  $\exp(\pm im\theta)$  symmetry where  $m = 1$ . We have seen that the amplitude of the magnetic pressure force correlates with the amplitude of the resonant Alfvén wave, so the resonant Alfvén wave will also have an  $m = 1$  symmetry (as in normal-mode analysis). It is therefore inevitable that the kink wave coexists with an  $m = 1$  Alfvén wave that oscillates at the kink frequency (Ruderman and Roberts, 2002; Terradas et al., 2006).

Moving to more general loop equilibria, such as the elliptical cross-section considered by Ruderman (2003) or the multi-stranded loop of Terradas et al. (2008), field-line resonance continues to imprint the spatial form of the global mode of oscillation on localised Alfvénic motions. This explains the complex variation of energy-density on the resonant surface previously seen in Terradas et al. (2008) (their Figure 10). The amplitude of the global (kink-like) mode of oscillation must vary around this surface, with the variation of the magnetic pressure gradient captured in the energy-density of the Alfvén wave.

In the analytic solution, magnetic pressure is not the only contribution to amplitude variations of the resonant Alfvén wave. Referring to equation (2.175), the Alfvénic velocity perturbation has a  $Y$  dependence of the form

$$\frac{1}{[\partial\rho/\partial X]_{X=0}} \left( \frac{1}{h_{Y0}} \left[ \frac{\partial b_z}{\partial Y} \right]_{X=0} \right),$$

where  $\partial\rho/\partial X \equiv |\nabla\rho|$ . When analysing numerical results, multiplying the magnetic pressure gradient by the inverse of the density gradient did not significantly improve correlation with the Alfvénic velocity perturbation. It is likely that no improvement was seen because the normal-mode result corresponds to the limit  $t \rightarrow \infty$ , and is therefore approximate within the runtime of our simulations.

The challenge of obtaining a complete analytic solution for three dimensional field-line resonance with a straight, uniform magnetic field is now ready for completion. Schulze-Berge et al. (1992) provide a framework, density variation along field-lines may be handled by the methods of Thompson and Wright (1993) and two dimensional variation of field-line Alfvén eigenfrequencies can be treated using the methods of this chapter. This would be a worthwhile project, and a keystone in the theory of field-line resonance. My expectation, however, is that by splitting the problem into its underlying parts (considered separately by Thompson and Wright (1993) and this present work) we have already obtained the key features of that ‘complete’ solution.

# Self-Consistent Treatment of Magnetosphere-Ionosphere Coupling Via Alfvén Waves

---

## 3.1 Introduction

This chapter sets out the methods – the model, governing equations and computer codes – used to investigate self-consistent magnetosphere-ionosphere (MI) coupling in later chapters. We have in fact developed two complementary approaches, which are set out side by side. The first of these uses ideal MHD to describe the magnetosphere and is based on the work of Cran-McGreehin et al. (2007). There are situations, however, where this model breaks down, due to formation of discontinuities by MI-coupling. This motivates the inclusion of electron inertial effects in the magnetosphere, for which we develop a second approach, guided in part by Lysak and Song (2008). The ideal approach is a very valuable tool: being simpler than the inertial model, it produces powerful analytic results and provides significant insight into the most fundamental behaviour of the system. On the other hand, the inertial model opens new areas of physics and new behaviours. As we shall see, it is by combining both approaches that we are able to achieve our primary aim of explaining the origin of small-scale Alfvén waves in large-scale current systems.

## 3.2 Model

We use a 2D model in which a thin ‘sheet’ E-region separates Earth’s atmosphere and magnetosphere (illustrated in Figure 3.1); assuming, for greater simplicity, that thermal plasma pressure plays a negligible role in the magnetosphere, and that Earth’s equilibrium magnetic field passes vertically through the E-region (a reasonable approximation at high magnetic latitudes). Cartesian coordinates  $(x, y, z)$  lend themselves to this model: we use  $z$  for ‘up’,  $x$  for ‘East–West’ and  $y$  for ‘North–South’. Driving comes from a disturbance in the outer magnetosphere, whose influence is included as a downward-propagating Alfvén wave (grossly exaggerated in Figure 3.1) with its magnetic perturbation aligned in the  $x$ -direction. Two further simplifications are made: first, we

decouple magnetospheric Alfvén waves from other MHD wave modes by making the model invariant in  $x$ ; and second, we assume that the incident Alfvén wave from the outer magnetosphere remains constant over the time-scale of interest. An identical setup was used by Cran-McGreehin et al. (2007).

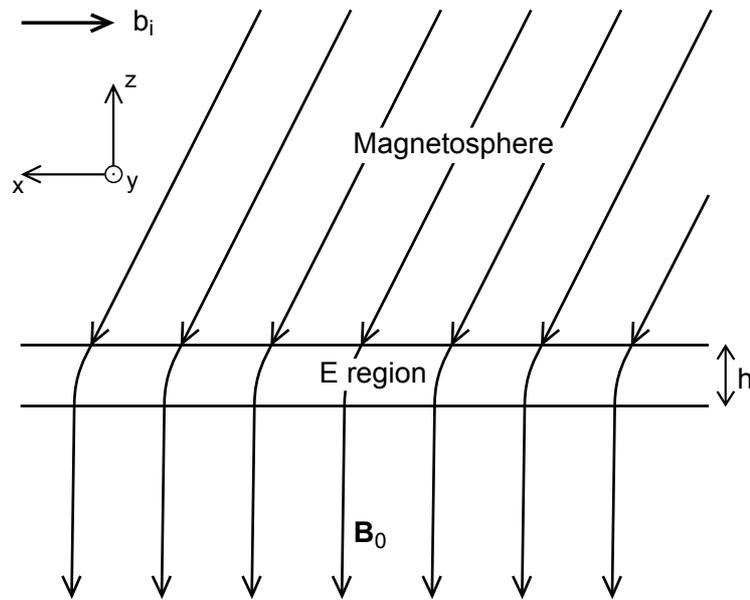


Figure 3.1: Diagram of the  $x$ - $z$  plane of the model, after Cran-McGreehin et al. (2007). An incident Alfvén wave is partially reflected from a sheet ionosphere. Copyright 2007 American Geophysical Union. Modified by permission of American Geophysical Union.

If the incident Alfvén wave varies in  $y$ , then it shears the magnetic field, resulting in field-aligned currents (FACs). Where these currents meet the E-region, they close through it as Pedersen current (the underlying atmosphere is a perfect insulator so currents cannot close there). Figure 3.2 illustrates such a current system.

Electrical currents rapidly establish a feedback, by which the magnetosphere and E-region influence one another. In the magnetosphere, field-aligned currents are carried by movement of electrons along the background magnetic field, and where these currents meet the E-region, they modify E-region electron density: downward field-aligned currents remove electrons from the E-region, locally decreasing electron number density; and upward field-aligned currents deposit electrons into the E-region (possibly causing additional ionisation if the deposited electrons carry sufficient energy), leading to a local increase in electron number density. Meanwhile, quasi-neutrality is preserved in the E-region by movement of ions as Pedersen current, which, in turn, induces electromagnetic fields in the magnetosphere. Hence, a back reaction occurs, modifying the original magnetospheric currents and completing the loop of ionosphere-magnetosphere coupling.

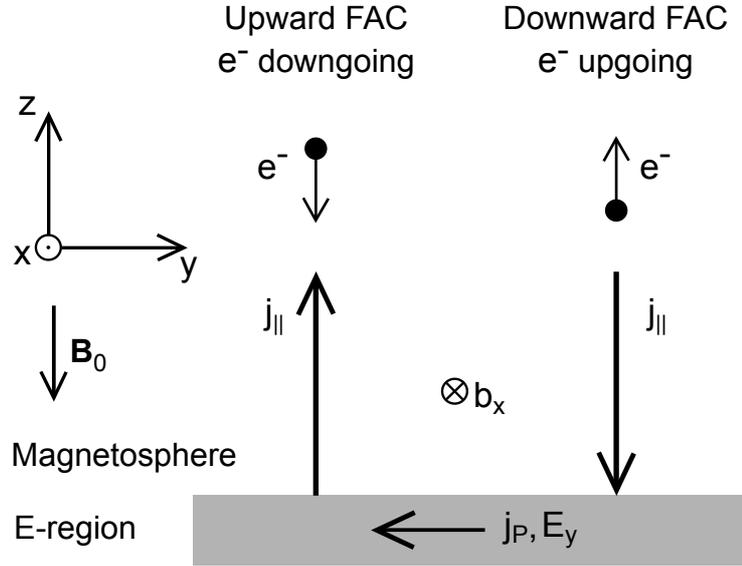


Figure 3.2: Diagram of the y-z plane of the model. A general Alfvén wave, incident on a sheet ionosphere, produces channels of upward and downward field-aligned current (FAC) in the magnetosphere, which close via Pedersen current in the E-region. Upward FAC deposits electrons into the E-region, locally enhancing number densities, whilst downward FAC removes electrons, causing depletion.

### 3.3 Governing Equations

#### 3.3.1 Coupling of Magnetosphere and Ionosphere

Two equations are required to capture the physics of MI-coupling. One of these represents the alteration of E-region number density due to field-aligned currents and the other embodies the effect of ionospheric Pedersen current on electromagnetic fields in the magnetosphere.

E-region number density is governed by the electron continuity equation

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial j_z}{\partial z} + \alpha (n_e^2 - n^2), \quad (3.1)$$

where,  $n$  is the E-region electron number density,  $j_z$  is the vertical (field-aligned) current,  $e$  is the fundamental charge,  $\alpha$  is the recombination coefficient (which we take to be constant), and  $n_e$  is the equilibrium electron density in the absence of field-aligned currents (e.g. Cran-McGreehin et al. (2007)). Equation (3.1) can be integrated over the thickness of the E-region. If we make the simplifying assumption that  $n$  is independent of height, then the height-integrated electron density is  $N = hn$ , where  $h$  is the thickness of the E-region. Taking  $j_z = 0$  in the atmosphere, integration

of (3.1) gives

$$\frac{\partial N}{\partial t} - \frac{j_z}{e} = \frac{\alpha}{h} (N_e^2 - N^2), \quad (3.2)$$

where  $j_z$  now represents the current at the top of the E-region, i.e.  $j_z(y, z = h)$ .

Examining equation (3.2), one sees that that if  $j_z$  is positive at the top of the E-region (upward field-aligned current) then E-region number density increases in time. Conversely, if  $j_z$  is negative at the top of the E-region (downward field-aligned current) then E-region number density decreases with time. This is the anticipated effect by which the magnetosphere changes E-region number density.

Next we consider the impact of ionospheric Pedersen currents on the magnetosphere. The Pedersen current in the E-region obeys

$$j_y = \frac{1}{\mu_0} \frac{\partial b_x}{\partial z} = \sigma_P E_y, \quad (3.3)$$

where  $\sigma_P$  is the Pedersen conductance (inverse of resistivity). Since the atmosphere is an insulator, it does not permit electrical currents, hence

$$j_z = -\frac{1}{\mu_0} \frac{\partial b_x}{\partial y} = 0 \quad (3.4)$$

at the base of the ionosphere. Subsequently, if  $b_x \rightarrow 0$  at  $z = 0$ ,  $y \rightarrow \pm\infty$ , then  $b_x = 0$  everywhere at the base of the ionosphere. Integrating (3.3) over the thickness of the E-region, assuming  $E_y$  remains approximately constant over the height of the (thin) E-region, and using  $b_x = 0$  at the base of the ionosphere, we find

$$b_x = \mu_0 \Sigma_P E_y \quad (3.5)$$

where  $b_x$  and  $E_y$  are now evaluated at the top of the E-region. Here  $\Sigma_P$  is the height-integrated Pedersen conductance, which is proportional to  $N$ . The dependence on  $N$  can be made explicit if we define  $\Sigma_{P0}$  as the height-integrated Pedersen conductance when  $N = N_e$ , and write

$$b_x = \frac{\mu_0 \Sigma_{P0}}{N_e} N E_y. \quad (3.6)$$

### 3.3.2 Magnetospheric Equations

A cold, collisionless plasma is described by

$$\text{Faraday's law:} \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \quad (3.7)$$

$$\text{Momentum equation: } \rho \frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u} \cdot \nabla) \mathbf{u} = \mathbf{j} \times \mathbf{B}, \quad (3.8)$$

$$\text{Generalised Ohm's law: } \mathbf{E} + \mathbf{u} \times \mathbf{B} = \frac{m^+ m^-}{\rho e^2} \frac{\partial \mathbf{j}}{\partial t} + \frac{m^+}{\rho e} \mathbf{j} \times \mathbf{B}, \quad (3.9)$$

$$\text{Ampère's law: } \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{B} - \mu_0 \mathbf{j}; \quad (3.10)$$

where  $\mathbf{E}$  and  $\mathbf{B}$  are electric and magnetic fields,  $\mathbf{u}$  is the plasma velocity,  $\rho$  and  $\mathbf{j}$  are the plasma and current densities, and  $m^+$  and  $m^-$  are the ion and electron masses. Here, we will obtain a set of four model equations by linearising the above and selecting the appropriate components for an Alfvén wave (polarised so that the magnetic field perturbation is aligned with the unit vector  $\hat{\mathbf{x}}$ ).

The  $x$ -components of Faraday's law and the momentum equation give

$$\frac{\partial b_x}{\partial t} = \frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}, \quad (3.11)$$

$$\rho_0 \frac{\partial u_x}{\partial t} = B_0 j_y. \quad (3.12)$$

Meanwhile, the parallel component of Ohm's law gives

$$\frac{\partial j_z}{\partial t} = \frac{E_z}{\mu_0 \lambda_e^2}, \quad (3.13)$$

where

$$\lambda_e = \sqrt{\frac{m^- m^+}{\mu_0 \rho_0 e^2}} \equiv \sqrt{\frac{m^-}{\mu_0 n_m e^2}} \quad (3.14)$$

is the electron inertial length, a length-scale obtained from the background plasma-density.

As it stands, the perpendicular component of Ohm's law,

$$\mathbf{E}_\perp + \mathbf{u} \times \mathbf{B} = \frac{m^+ m^-}{\rho e^2} \frac{\partial \mathbf{j}_\perp}{\partial t} + \frac{m^+}{\rho e} \mathbf{j} \times \mathbf{B}, \quad (3.15)$$

contains a convective term on the left-hand side, and inertial and Hall terms on the right-hand side. If we restrict ourselves to MHD time-scales, then the electric fields produced by Hall and inertial effects are much smaller than the electric field produced by convection, and can be neglected with little error (e.g. Boyd and Sanderson (1969)). This is a substantial simplification, so we will use

$$E_y = -B_0 u_x. \quad (3.16)$$

The inertial term is retained in the parallel component of Ohm's law (3.13) because it is the only source of parallel electric field. Generally,  $E_z \ll E_y$ , but equation (3.11) shows that  $E_z$  can nonetheless play a significant role in Alfvén wave dynamics if gradients in  $y$  are much steeper than gradients in  $z$ . We therefore retain finite  $E_z$  in our model.

The remaining equations come from Ampère's law. First, the parallel component states

$$\frac{\partial E_z}{\partial t} = -c^2 \left( \mu_0 j_z + \frac{\partial b_x}{\partial y} \right). \quad (3.17)$$

Similarly, the  $y$ -component yields

$$\frac{1}{c^2} \frac{\partial E_y}{\partial t} = -\mu_0 j_y + \frac{\partial b_x}{\partial z}, \quad (3.18)$$

which can be simplified if we notice that equations (3.12) and (3.16) can be combined to give

$$\mu_0 j_y = \frac{1}{v_A^2} \frac{\partial E_y}{\partial t}, \quad (3.19)$$

where  $v_A = B_0/\sqrt{\mu_0 \rho_0}$  is the background Alfvén speed. This equation reveals that  $j_y$  in the magnetosphere is carried by a polarisation drift of ions. Using (3.19) to eliminate  $j_y$  from (3.18),

$$\frac{1}{c^2} \frac{\partial E_y}{\partial t} = -\frac{1}{v_A^2} \frac{\partial E_y}{\partial t} + \frac{\partial b_x}{\partial z} \quad (3.20)$$

$$\Rightarrow \frac{\partial E_y}{\partial t} = \frac{v_A^2}{1 + v_A^2/c^2} \frac{\partial b_x}{\partial z}. \quad (3.21)$$

Bringing the above together, our model is characterised by four wave fields,  $(b_x, E_y, E_z, j_z)$ , and these satisfy

$$\frac{\partial b_x}{\partial t} = \frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}, \quad (3.22)$$

$$\frac{\partial E_y}{\partial t} = \frac{v_A^2}{1 + v_A^2/c^2} \frac{\partial b_x}{\partial z}, \quad (3.23)$$

$$\frac{\partial E_z}{\partial t} = -c^2 \left( \mu_0 j_z + \frac{\partial b_x}{\partial y} \right), \quad (3.24)$$

$$\frac{\partial j_z}{\partial t} = \frac{E_z}{\mu_0 \lambda_e^2}. \quad (3.25)$$

These model equations automatically satisfy  $\nabla \cdot \mathbf{b} = 0$ , since  $x$  is the invariant direction, and are in fact a limiting case of the governing equations used by Lysak and Song (2008) to model Alfvén wave propagation inside the ionospheric Alfvén resonator.

The physical value of  $c$  is  $3 \times 10^8 \text{ ms}^{-1}$ , so terms in  $1/c^2$  can typically be neglected, as we show here. In situations of strong magnetic field and low plasma-density  $v_A$  may become as large as a tenth of the speed of light, but even then  $v_A^2 \ll c^2$ , so the ‘relativistic Alfvén speed’,

$$v_{A,rel} = \frac{v_A}{\sqrt{1 + v_A^2/c^2}}, \quad (3.26)$$

which appears in (3.23), reduces to the non-relativistic Alfvén speed,  $v_A$ , with little error. Similarly, the left-hand side of (3.24) can be neglected for low-frequency solutions, provided that the angular frequency of oscillations,  $\omega$ , satisfies

$$\omega^2 \ll \omega_{pe}^2, \quad (3.27)$$

where

$$\omega_{pe} = \frac{c}{\lambda_e} \quad (3.28)$$

is the electron plasma frequency.

The speed of light enters equations (3.23) and (3.24) through the displacement current. If the displacement current in (3.24) had been non-negligible, then the time evolution of all magnetospheric fields could have been solved explicitly. This suggests a useful slight of hand that is exploited in Section 3.5, despite the large physical value of  $c$ .

For analytic work, we neglect the displacement current by the arguments above. This immediately produces the familiar component of Ampère's law,

$$j_z = -\frac{1}{\mu_0} \frac{\partial b_x}{\partial y}, \quad (3.29)$$

that determines  $j_z$  from  $b_x$ . Since equation (3.25) determines  $E_z$  from the rate of change of  $j_z$ , the system may be written in terms of just two wave fields,  $E_y$  and  $b_x$ , that satisfy

$$\frac{\partial b_x}{\partial t} - \frac{\partial}{\partial y} \left( \lambda_e^2 \frac{\partial^2 b_x}{\partial y \partial t} \right) = \frac{\partial E_y}{\partial z}, \quad (3.30)$$

$$\frac{\partial E_y}{\partial t} = v_A^2 \frac{\partial b_x}{\partial z}. \quad (3.31)$$

These are the MHD equations for an inertial Alfvén wave.

An important limit is found if length scales perpendicular to the background magnetic field are much larger than the electron inertial length, i.e.  $\lambda_y \gg \lambda_e$ . In this limit, the second term on the left-hand side of (3.30) is negligible, and finite electron mass plays a vanishing role in the magnetosphere. Now electrons respond instantly to parallel electric field, so that  $E_z \approx 0$  at all times. Under such conditions, both the magnetosphere and the Alfvén wave are described as *ideal*, the MHD equations for an ideal Alfvén wave being

$$\frac{\partial b_x}{\partial t} = \frac{\partial E_y}{\partial z}, \quad (3.32)$$

$$\frac{\partial E_y}{\partial t} = v_A^2 \frac{\partial b_x}{\partial z}. \quad (3.33)$$

Later, we shall wish to discuss the dispersion relation for Alfvén waves. This is obtained by assuming wave fields to be normal-modes, varying as  $\exp(i(\mathbf{k} \cdot \mathbf{r} - \omega t))$ . Then, equations (3.30) and (3.31) can be combined to give

$$\omega = \frac{k_z v_A}{\sqrt{1 + k_y^2 \lambda_e^2}}. \quad (3.34)$$

In the ideal limit,  $k_y^2 \lambda_e^2 \ll 1$ , so the dispersion relation simply becomes

$$\omega = k_z v_A. \quad (3.35)$$

### 3.3.3 Single Equation for Ionosphere Coupled to an Ideal Magnetosphere

Inspection of equations (3.32) and (3.33) reveals two special cases of ideal Alfvén wave that are of particular interest. If  $v_A$  is independent of  $z$ , then setting

$$E_y = \pm v_A b_x \quad (3.36)$$

gives

$$\frac{\partial f}{\partial t} \mp v_A \frac{\partial f}{\partial z} = 0, \quad (3.37)$$

where  $f$  can be either  $E_y$  or  $b_x$ . Equation (3.37) is a special type of advection equation that describes waves travelling along  $z$  at a speed  $\mp v_A$ , retaining the shape of the disturbance. The direction of propagation is easily confirmed from the sign of the  $z$ -component of the Poynting vector,  $S_z = -E_y b_x / \mu_0$ , so, for ideal Alfvén waves incident on the ionosphere (subscript  $i$ ) and reflected from the ionosphere (subscript  $r$ ),

$$E_i = v_A b_i, \quad (3.38)$$

$$E_r = -v_A b_r. \quad (3.39)$$

The ability to split an ideal Alfvén wave into its constituent parts, one upgoing, one downgoing and both non-dispersive, allows the combination of (3.38) and (3.39) with the coupling equations (3.2) and (3.6) to obtain a single governing equation (Cran-McGreehin et al., 2007). We start by writing the fields in (3.6) as a sum of incident and reflected ideal Alfvén waves:

$$(b_i + b_r) = \mu_0 \Sigma_{P0} \frac{N}{N_e} (E_i + E_r). \quad (3.40)$$

Using equations (3.38) and (3.39) to eliminate magnetic perturbations, then doing a little algebra,

$$\left(1 - \frac{E_r}{E_i}\right) = \beta \frac{N}{N_e} \left(1 + \frac{E_r}{E_i}\right) \quad (3.41)$$

where

$$\beta = \mu_0 v_A \Sigma_{P0} \quad (3.42)$$

is the ratio of the equilibrium height-integrated Pedersen conductance to the ideal Alfvén conductance  $\Sigma_A = 1/\mu_0 v_A$ . Rearranging for  $E_r/E_i$ , and defining this quantity as the ionospheric reflection coefficient,  $r$ ,

$$r = \frac{E_r}{E_i} = \frac{1 - \beta N/N_e}{1 + \beta N/N_e}. \quad (3.43)$$

The reflection coefficient also determines the reflected magnetic field, since

$$\frac{b_r}{b_i} = -\frac{E_r}{E_i} = -r. \quad (3.44)$$

Thus, the total magnetic field perturbation at the top of the E-region is

$$b_x = (1 - r)b_i = \frac{2\beta N b_i}{N_e + \beta N}. \quad (3.45)$$

Using (3.45) and the  $z$ -component of Ampère's law stated in (3.29), the height-integrated ionospheric continuity equation (3.2) becomes

$$\frac{\partial N}{\partial t} + \frac{1}{\mu_0 e} \frac{\partial}{\partial y} \left( \frac{2\beta N b_i}{N_e + \beta N} \right) = \frac{\alpha}{h} (N_e^2 - N^2). \quad (3.46)$$

Equation (3.46) is a first-order partial differential equation for  $N$ , that completely characterises the time evolution of an ionosphere coupled to an ideal magnetosphere. The incident Alfvén wave is specified through  $b_i(y)$ , which is independent of time in our model, and  $N$  is solved using (3.46). All magnetospheric fields follow through the reflection coefficient and the advective property of incident and reflected ideal Alfvén waves.

In previous treatments, this same analysis has been presented using the incident velocity perturbation

$$\mathbf{u}_i = \frac{1}{B_0^2} (\mathbf{E}_i \times \mathbf{B}_0) \quad (3.47)$$

in place of  $\mathbf{b}_i$ . This is true of Cran-McGreehin et al. (2007) and Russell et al. (2010). In the

Northern Hemisphere, defining the velocity perturbation as  $-u_i \hat{\mathbf{x}}$  gives

$$\begin{aligned} E_i &= v_A b_i = B_0 u_i, \\ \Rightarrow b_i &= \frac{B_0 u_i}{v_A} \end{aligned} \quad (3.48)$$

and recovers the result in those papers. We will, however, use the magnetic perturbation for two reasons: (i) it produces a governing equation that is independent of hemisphere; and (ii) it is more consistent with the equations used to include electron inertia in the magnetosphere.

### 3.4 Numerical Code with Ideal Magnetosphere

A theme that becomes familiar in this thesis is the use of computer codes to gain insight into complicated physical systems, the codes providing an important opportunity to explore models in a hands-on way. This section details a 1D code that solves an equation equivalent to (3.46), and presents tests used to verify its proper working. Results from the code are presented in Chapters 4, 5 and 6.

#### 3.4.1 1D Governing Equation in Form Solved by Code

It is convenient to have a computer solve governing equation (3.46) in dimensionless form. When we developed a code to do this, we chose to establish it for the Northern Hemisphere and specify the incident wave through  $u_i$  instead of  $b_i$ . At the time, this decision was made for consistency with Cran-McGreehin et al. (2007), but the analysis of results will be phrased in terms of  $b_i$  for reasons that have since come to light and are outlined at the end of Section 3.3.3.

Putting  $b_i = B_0 u_i / v_A$  and writing

$$\tilde{N} = \frac{N}{N_e}, \quad (3.49)$$

equation (3.46) may be written as

$$\frac{\partial \tilde{N}}{\partial t} + \frac{2}{\mu_0 e N_e} \frac{\partial}{\partial y} \left( \frac{\beta \tilde{N} B_0 u_i}{v_A (1 + \beta \tilde{N})} \right) = \frac{\alpha N_e}{h} (1 - \tilde{N}^2). \quad (3.50)$$

Distances are normalised with respect to a typical scale in  $y$ ,  $l_0$ , and velocities are normalised by the greatest amplitude of the incident Alfvén wave,  $u_{i0}$ . This gives rise to a typical time scale,  $\tau = l_0 / u_{i0}$ , that is used to normalise times. We therefore define

$$\tilde{u}_i = \frac{u_i}{u_{i0}}, \quad (3.51)$$

$$\tilde{y} = \frac{y}{y_0}, \quad (3.52)$$

$$\tilde{t} = \frac{t}{\tau} = \frac{u_{i0}}{l_0} t. \quad (3.53)$$

Making these normalisations, and introducing

$$\eta = \frac{2B_0 \Sigma_{P0}}{eN_e}, \quad (3.54)$$

$$\tilde{\alpha} = \frac{\alpha N_e l_0}{h u_{i0}}, \quad (3.55)$$

a little rearranging puts the governing equation into the form solved by the numerical code (also that employed in Cran-McGreehin et al. (2007) and Russell et al. (2010)):

$$\frac{\partial \tilde{N}}{\partial \tilde{t}} + \frac{\partial}{\partial \tilde{y}} \left( \frac{\eta \tilde{u}_i \tilde{N}}{1 + \beta \tilde{N}} \right) = \tilde{\alpha} (1 - \tilde{N}^2). \quad (3.56)$$

### 3.4.2 Numerical Methods

The computer code takes an initial condition for  $\tilde{N}$  at  $\tilde{t} = 0$ , and evolves it in time under the influence of an incident Alfvén wave specified through  $\tilde{u}_i(\tilde{y})$ . Since different combinations of incident waves and initial conditions are used in later chapters, we leave it until then to detail specifics, but do require that  $\tilde{u}_i$  be negative in the solution domain, and  $\tilde{u}_i = d\tilde{u}_i/d\tilde{y} = 0$  at boundaries in  $\tilde{y}$ .

As we shall see, discontinuities in  $\tilde{N}$  naturally form in an ionosphere coupled to an ideal magnetosphere, then propagate in the direction of the incident electric field which is negative for negative  $\tilde{u}_i$ . Propagating discontinuities lead to numerical instability if spatial derivatives are represented as centred finite differences, but the problem can be avoided by ‘upwinding’, that is to say, by using a one-sided finite difference that propagates information in the same direction that the discontinuity travels. We therefore evaluate spatial derivatives at the  $j^{\text{th}}$  grid-point using

$$\left( \frac{\partial \tilde{f}}{\partial \tilde{y}} \right)_j \equiv \frac{\tilde{f}_{j+1} - \tilde{f}_j}{\Delta \tilde{y}}, \quad (3.57)$$

where  $\Delta \tilde{y}$  is the grid-spacing in  $y$ .

The code uses an Euler scheme to advance in time, and combined with our choice of finite difference for spatial derivatives, this gives a code that is first-order accurate in both space and time. These low orders are actually an advantage, due to the multi-scale nature of magnetosphere-ionosphere coupling. In simulations that form small scales, it is desirable to use a large number of grid-points to resolve these scales, and the Courant-Friedrichs-Lewy stability condition requires

a correspondingly short time-step. As an example, this code typically uses a grid-spacing  $\Delta\tilde{y} = 0.001$  and a time-step  $\Delta\tilde{t} = 0.0001$ . Because of the large number of grid-points and time-steps, a higher order scheme would not noticeably alter the results, but would increase runtime. First order, therefore, provides a good balance of accuracy, stability and efficiency.

### 3.4.3 Testing

Before using a computer code to produce new results, it is important to check it for human errors that may have crept in during development (e.g. through bad design or programming), and also to establish its accuracy after numerical errors (e.g. accumulation of round-off errors). With this in mind, the code was put through a testing phase.

The first test that we developed is conservation of total electron number for  $\tilde{\alpha} = 0$ . If ionisation and recombination are neglected in the E-region, then equation (3.56) may be integrated with respect to  $y$  over the simulation domain to give

$$\frac{\partial}{\partial\tilde{t}} \int_{\text{domain}} \tilde{N} d\tilde{y} = \left[ -\frac{\eta\tilde{u}_i\tilde{N}}{1+\beta\tilde{N}} \right]_{\text{left edge of domain}}^{\text{right edge of domain}} = 0. \quad (3.58)$$

The final equality follows from our assumption that  $\tilde{u}_i = 0$  at the edges of the domain, meaning that the Alfvén wave is localised in  $\tilde{y}$  and does not extend beyond our domain. This also has the property that upward and downward field-aligned currents are balanced at the top of the E-region, and hence there is no net flow of electrons into the E-region. Therefore, the total number of E-region electrons,  $\int_{\text{domain}} \tilde{N} d\tilde{y}$  is conserved if  $\tilde{\alpha} = 0$  and  $\tilde{u}_i = 0$  at the edges of the simulation domain.

This was tested using the profile of  $\tilde{u}_i$  detailed in Chapter 4, with  $\beta = 100$ ,  $\eta = 1.015$ ,  $\delta\tilde{y} = 0.001$ ,  $\delta\tilde{t} = 0.0001$  and  $\tilde{\alpha} = 0$ . During the test, the total electron number, evaluated over the numerical domain in dimensionless form, varied from its initial value of 8.0 by less than  $1.78 \times 10^{-14}$ . This is consistent with errors from numerical round-off, indicating that the code satisfactorily conserves total electron number density if  $\tilde{\alpha} = 0$ .

To further provide confidence in the code, a version was modified to reproduce numerical results from Cran-McGreehin et al. (2007). This required a symmetric boundary at  $\tilde{y} = 0$  and a change to ‘left-looking’ finite differences for derivatives in  $\tilde{y}$ . With these alterations, the code successfully reproduced Figure 3 of Cran-McGreehin et al. (2007) without any anomalies.

### 3.4.4 Convergence to Steady State

The governing equation (3.56) provides a test of convergence to the steady state. Taking  $\partial\tilde{N}/\partial\tilde{t} = 0$  and  $\tilde{\alpha} \neq 0$ , integration of (3.56) over  $\tilde{y}$  gives

$$\int_{-\infty}^{\infty} (1 - \tilde{N}^2) d\tilde{y} = 0, \quad (3.59)$$

showing that ionisation and recombination balance in the steady state. The integral quantifies convergence to the steady state: for example, the numerical steady states presented in Chapter 4 were obtained by running the code until  $\left| \int (1 - \tilde{N}^2) d\tilde{y} \right|$  was less than  $10^{-10}$  and converging to zero, as evaluated over the simulation domain.

## 3.5 Numerical Code with Electron Inertial Effects

### 3.5.1 Strategy for Magnetospheric Fields

We want to solve equations (3.22) to (3.25) in the limit where  $c^2 \gg v_A^2$  and the angular frequency is much less than the electron plasma frequency defined in (3.28). In designing a code to obtain numerical solutions, two distinct approaches are available.

The first approach is to take the limit analytically, then solve equations (3.30) and (3.31) numerically. This has the advantage that the limit is guaranteed, and is the option chosen by Wei et al. (1994); on the other hand, we must also contend with the fact that equation (3.30) provides  $b_x$  at a new time level implicitly, requiring the use of special implicit numerical techniques. Alternatively, it is equally possible to solve (3.22) to (3.25) explicitly by taking the limit ‘approximately’. Presented with these two options, we preferred to use the latter, enabling us to apply our prior experience with explicit methods.

The way to an explicit solution is pointed to by Lysak and Song (2008), with reasoning as follows. Taking the limit  $v_A^2 \ll c^2$  in (3.23) presents no problem, so we do so, and let it become equation (3.31). All three of (3.22), (3.25) and (3.31) are written in explicit form and their numerical solution presents no difficulties. The problem lies entirely in the solution of equation (3.24). In principle, (3.24) allows explicit solution for  $E_z$  at a new time level, but numerical stability requires an incredibly short time-step because  $c = 3 \times 10^8 \text{ ms}^{-1}$ . The time-step restriction is equivalent to resolving electron plasma oscillations, which can be seen as a natural part of the solution by eliminating  $E_z$  between (3.24) and (3.25) to give

$$\frac{\partial^2 j_z}{\partial t^2} + \omega_{pe}^2 j_z = -\frac{\omega_{pe}^2}{\mu_0} \frac{\partial b_x}{\partial y}. \quad (3.60)$$

Here,  $\omega_{pe}$  is the electron plasma frequency defined in (3.28). The solution to the homogeneous version of (3.60) oscillates at the electron plasma frequency, and these oscillations remain a feature of the inhomogeneous equation too (this is discussed again in Section 3.5.3.3).

Fortunately, there is a numerical sleight-of-hand that produces accurate solutions at a longer time-step by artificially lowering the electron plasma frequency. The reasoning is as follows. Let us take equation (3.24) and write it down, replacing the universal constant  $c$  with  $c_{\parallel}$ , which we treat as a variable parameter:

$$\frac{\partial E_z}{\partial t} = -c_{\parallel}^2 \left( \mu_0 j_z + \frac{\partial b_x}{\partial y} \right). \quad (3.61)$$

If we put  $c_{\parallel} = c$ , then  $\mu_0 j_z$  and  $-\partial b_x / \partial y$  differ by about one part in  $10^{17}$ . A physically accurate solution, therefore, has  $\mu_0 j_z \approx -\partial b_x / \partial y$ . In pursuit of a numerical solution, however, we might be quite satisfied to set  $c_{\parallel} = 100$ , so that  $\mu_0 j_z$  and  $-\partial b_x / \partial y$  differ by about one part in  $10^4$ . Sufficient accuracy is obtained, but the number of time-steps required for a stable solution is drastically reduced.

As well as allowing a longer time-step, artificial reduction of  $c_{\parallel}$  also alleviates problems caused by round-off error. Double-precision (64-bit) arithmetic is accurate to about one part in  $10^{19}$ . If  $\mu_0 j_z$  and  $-\partial b_x / \partial y$  differ by about one part in  $10^{17}$  (as for  $c_{\parallel} = c$ ) then round-off errors lead to a 1% error in each update of  $E_z$ , which is very undesirable. However, if  $c_{\parallel} = 100$  so that  $\mu_0 j_z$  and  $-\partial b_x / \partial y$  differ by about one part in  $10^4$ , the significance of round-off error in equation (3.61) becomes much more tolerable, to the point that we need not show it any concern.

In order to decide a suitable value for  $c_{\parallel}$ , several runs were performed. These runs were stable and demonstrate the expected convergence as  $c_{\parallel}$  becomes large. They are detailed in Section 3.5.5, where we also discuss the final choice of  $c_{\parallel}$ .

### 3.5.2 Magnetospheric Equations in Form Solved by Code

It is convenient to have a computer solve governing equations in dimensionless form. This time, we normalise distances with respect to  $l_0$ , a typical distance in  $y$ , and velocities with respect to  $v_{A0}$ , the greatest background Alfvén speed. These can be combined to give a typical time scale,  $\tau = l_0 / v_{A0}$ , which we use to normalise times. We therefore put

$$\tilde{\mathbf{r}} = \frac{\mathbf{r}}{l_0}, \quad (3.62)$$

$$\tilde{\mathbf{v}} = \frac{\mathbf{v}}{v_{A0}}, \quad (3.63)$$

$$\tilde{t} = \frac{t}{\tau} = \frac{v_{A0}}{l_0} t. \quad (3.64)$$

Introducing  $\epsilon$  as a dimensionless scaling parameter, Alfvén wave fields are scaled as

$$\tilde{\mathbf{b}} = \frac{\mathbf{b}}{\epsilon B_0}, \quad (3.65)$$

$$\tilde{\mathbf{E}} = \frac{\mathbf{E}}{\epsilon B_0 v_{A0}}, \quad (3.66)$$

$$\tilde{\mathbf{j}} = \frac{l_0 \mu_0}{\epsilon B_0} \mathbf{j}. \quad (3.67)$$

Using these normalisations, equations (3.22), (3.25), (3.31) and (3.61) become

$$\frac{\partial \tilde{b}_x}{\partial \tilde{t}} = \frac{\partial \tilde{E}_y}{\partial \tilde{z}} - \frac{\partial \tilde{E}_z}{\partial \tilde{y}}, \quad (3.68)$$

$$\frac{\partial \tilde{E}_y}{\partial \tilde{t}} = \tilde{v}_A^2 \frac{\partial \tilde{b}_x}{\partial \tilde{z}}, \quad (3.69)$$

$$\frac{\partial \tilde{E}_z}{\partial \tilde{t}} = -\tilde{c}_{\parallel}^2 \left( \tilde{j}_z + \frac{\partial \tilde{b}_x}{\partial \tilde{y}} \right), \quad (3.70)$$

$$\frac{\partial \tilde{j}_z}{\partial \tilde{t}} = \frac{\tilde{E}_z}{\tilde{\lambda}_e^2}, \quad (3.71)$$

which are the magnetospheric equations to be solved by computer.

Equations (3.68) to (3.71) are linear, and so have arbitrary amplitude. We therefore use a dimensionless scaling parameter,  $\epsilon$ , to scale the amplitude of magnetospheric perturbations to the order of unity, which is numerically desirable. The amplitude of  $j_z$  is not arbitrary when it comes to coupling to the non-linear E-region: as we shall see in Section 3.5.7, this leads to  $\epsilon$  becoming a simulation parameter.

### 3.5.3 Numerical Methods

#### 3.5.3.1 Scheme

Wave fields were advanced in time using the leapfrog-trapezoidal method detailed in Rickard and Wright (1994), and all spatial derivatives were evaluated as centred finite differences (a suitable choice because electron inertia prevents the formation of unresolved discontinuities). These choices make simulations second-order accurate in both space and time.

#### 3.5.3.2 Grid

Examining the spatial derivatives in equations (3.68) to (3.71), we have the option of staggering the grid  $\tilde{y}$  and/or  $\tilde{z}$ . In many computational problems, it is inefficient to solve for all fields on

a common grid: that approach is equivalent to simultaneously solving on multiple decoupled staggered grids, which is clearly redundant. Anticipating the ionospheric boundary-condition described in Section 3.5.7, the grid is staggered in  $\tilde{z}$ , but not in  $\tilde{y}$ . The staggered grid can be visualised if we think of grid-cells, each of which contains two grid-points (Figure 3.3). In each cell, the lower grid-point is described as a ‘whole’ grid-point because it has  $\tilde{z} = I\Delta\tilde{z}$  where  $I$  is an integer and  $\Delta\tilde{z}$  is the grid-spacing in  $\tilde{z}$  that separates cells. The upper point in the same cell is described as a ‘half’ grid-point, having  $\tilde{z} = (I + \frac{1}{2})\Delta\tilde{z}$ . The wave fields are then separated so that  $\tilde{b}_x$ ,  $\tilde{E}_z$  and  $\tilde{j}_z$  are defined at whole grid-points, and  $\tilde{E}_y$  is defined at half grid-points.

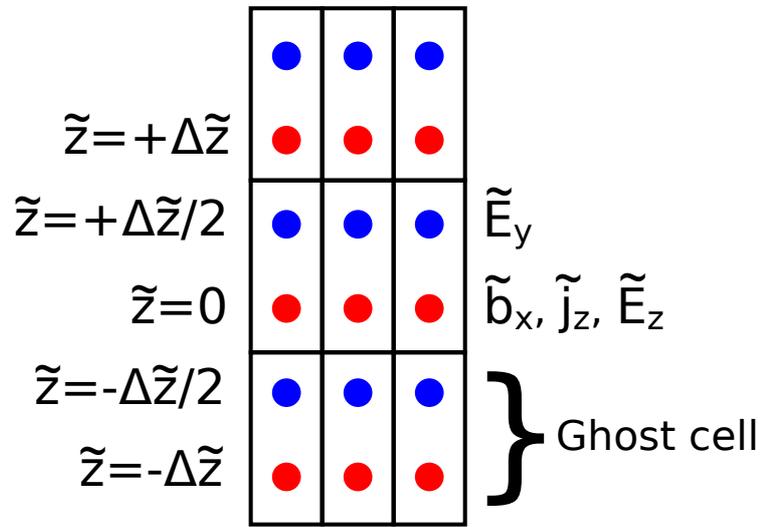


Figure 3.3: Fragment of the staggered grid used to solve for inertial Alfvén wave fields in the magnetosphere. Separation of  $\tilde{E}_y$  from the other magnetospheric fields removes inefficient redundancy, and also simplifies enforcement of the ionospheric boundary-condition at  $\tilde{z} = 0$ . A row of ghost-cells, used to enforce the bottom boundary-condition, is also shown.

Grid-spacing was chosen to balance resolution of small scales against efficiency. The smallest scales that develop are in the  $\tilde{y}$  direction (perpendicular to the background magnetic field) and these are of the order of the electron inertial length. With this in mind, the user is given a free choice for the number of grid-points in the  $\tilde{z}$  direction, but they are asked for the number of points with which to resolve an electron inertial length, which subsequently determines the number of grid-points in  $\tilde{y}$ .

### 3.5.3.3 Time-step

In order to obtain numerical stability, the time-step must resolve the shortest time-scale in the system, and the first consideration in this regard is resolving artificial plasma oscillations. The presence of these short period oscillations is revealed if we combine equations (3.70) and (3.71)

to give

$$\frac{\partial^2 \tilde{j}_z}{\partial \tilde{t}^2} + \tilde{\omega}_{pe}^2 \tilde{j}_z = -\tilde{\omega}_{pe}^2 \frac{\partial \tilde{b}_x}{\partial \tilde{y}}, \quad (3.72)$$

where

$$\tilde{\omega}_{pe} = \frac{\tilde{c}_{||}}{\tilde{\lambda}_e} \quad (3.73)$$

is the artificial plasma frequency. We can get a feel for the complete solution to (3.72) by writing perturbations as a sum of low and high frequency solutions. The low-frequency solution ( $\omega \ll \omega_e$ ) satisfies (3.72) to leading order by balancing the second and third terms to give  $\tilde{j}_z \approx -\partial \tilde{b}_x / \partial \tilde{y}$ . Meanwhile, it is evident that the higher-frequency solution will vary on a short time-scale of  $2\pi / \tilde{\omega}_{pe}$ . This leads us to expect that the complete solution for  $\tilde{j}_z$  will oscillate around  $\tilde{j}_z \approx -\partial \tilde{b}_x / \partial \tilde{y}$  at a frequency of approximately  $\tilde{\omega}_{pe}$ . Such a ‘multiple time-scales’ analysis, although approximate, indicates the likely presence of short-period oscillations, and shows that we should ensure the time-step resolves the artificial plasma period,

$$\tilde{\tau}_p = \frac{2\pi}{\tilde{\omega}_{pe}} = \frac{2\pi \tilde{\lambda}_e}{\tilde{c}_{||}}. \quad (3.74)$$

In addition to resolving artificial electron plasma oscillations, the time-step must satisfy the Courant-Friedrichs-Lewy condition, equivalent to resolving the shortest time that information takes to physically travel the distance between adjacent grid-cells. I have found it useful to first estimate this time from the grid-spacing and typical velocities, and then to use short trial simulations to fine tune the time-step for stability and efficiency.

### 3.5.4 Boundary-Conditions Using Ghost-Cells

Boundary-conditions were enforced using ghost-cells. This approach has won considerable favour with me, because, with a little care at the design stage, they contribute to flexible code that is easily modified to handle new situations. The principle is to surround the edge of the ‘physical’ domain with one or more layers of ‘ghosts’. Every time that a spatial derivative is required at a physical grid-point, it can, therefore, be computed from values at neighbouring cells using finite difference. In order to get the correct derivative at the edges of the physical domain, one simply ensures that appropriate values are assigned to the ghost-cells before the finite difference is computed.

An example clarifies the method. Say we wish to use ghost-cells to compute derivatives of a field  $\tilde{f}$ , requiring that  $\tilde{f}$  be symmetric at the left-hand edge of our domain. Second-order centred finite differences compute  $\partial \tilde{f} / \partial \tilde{y}$  at the left-hand edge of the domain using one grid-cell to the

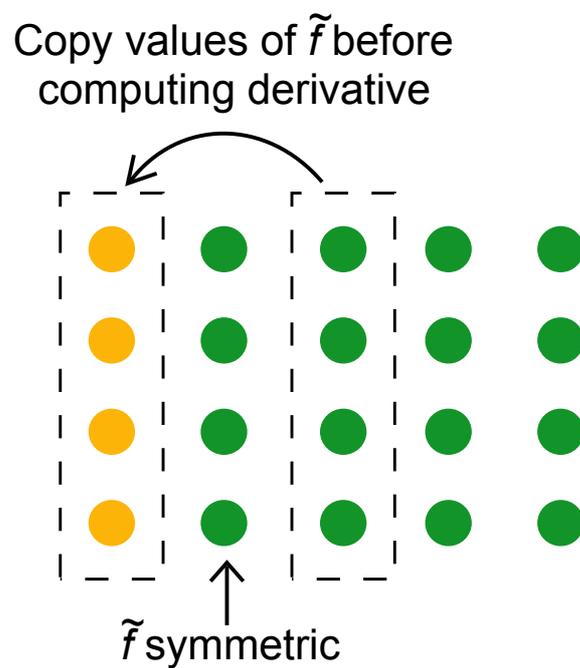


Figure 3.4: Use of ghost-cells to implement a symmetric boundary-condition. A layer of ghost-cells (yellow) is positioned along the left-hand edge of the physical domain (green), to provide values for computation of spatial derivatives. To make a field,  $f$ , symmetric at the left-hand edge of the domain, each ghost-cell is simply assigned the value of its mirror-image, before a finite difference is computed.

left and one grid-cell to the right. The grid-cell to the right exists, so it presents no problems; the grid-cell to the left, however, must be conjured, so a single layer of ghost-cells is positioned next to the left boundary. Figure 3.4 illustrates how the new cells are added. Now, in order to make  $\tilde{f}$  symmetric at the left-edge of the domain, we simply assign ghost-cells the value of their ‘mirror-image’ inside the physical domain. The beauty of this technique is that boundary-conditions are easily modified by changing the assignment step: no change need be made to the part of the code computing derivatives at physical cells.

### 3.5.5 Test Case: Normal-Mode Inertial Alfvén Wave

A number of tests were performed to establish a threshold for  $\tilde{c}_{||}$  beyond which results are physically realistic, and also to demonstrate proper working of the code. For the first test, the simulation was initialised at  $\tilde{t} = 0$  with  $\tilde{b}_x = \tilde{E}_z = \tilde{j}_z = 0$  everywhere and

$$\tilde{E}_y = \cos(\tilde{k}_y \tilde{y}) \sin(\tilde{k}_z \tilde{z}). \quad (3.75)$$

Boundary-conditions were enforced in keeping with a normal-mode solution: boundaries in  $\tilde{y}$  are nodes in  $\tilde{E}_z$  and  $\tilde{j}_z$ , and antinodes in  $\tilde{E}_y$  and  $\tilde{b}_x$ , whereas boundaries in  $\tilde{z}$  are nodes in  $\tilde{E}_y$ , and antinodes in  $\tilde{b}_x$ ,  $\tilde{E}_z$  and  $\tilde{j}_z$ . Background Alfvén speed and the electron inertial length were both uniform, and we experimented with the value of  $\tilde{c}_{||}$ .

In Section 3.5.3.3, artificial plasma oscillations were discussed, concluding that field-aligned current,  $\tilde{j}_z$ , is expected to oscillate about the MHD ( $\tilde{v}_A/\tilde{c}_{||} \rightarrow 0$ ) solution, with a period of approximately  $\tilde{\tau}_p = 2\pi\tilde{\lambda}_e/\tilde{c}_{||}$ . For a physically realistic solution (in the MHD regime), the plasma oscillations are negligible, which becomes increasingly true as  $\tilde{\tau}_p \rightarrow 0$ . Figure 3.5, shows  $\tilde{j}_z$  at a fixed point for a simulation with  $\tilde{\tau}_p = 0.2$  everywhere. The long period oscillation corresponds to code’s representation of MHD behaviour ( $\tilde{v}_A/\tilde{c}_{||} \rightarrow 0$ ), on top of which the artificial plasma oscillation is easily seen. Figure 3.6 is similar to 3.5, but is produced from a simulation with  $\tilde{\tau}_p = 0.05$  everywhere. As  $\tilde{\tau}_p$  is reduced, the period and amplitude of the artificial plasma oscillation both reduce, so that, this time, the plasma oscillation plays a minimal role and the MHD solution is dominant. We therefore set  $\tilde{c}_{||}$  in each simulation so that  $\tilde{\tau}_p \leq 0.05$  everywhere.

Our next test shows that the code produces correct normal-mode behaviour. If the code functions correctly, then a normal-mode should oscillate at the angular frequency predicted by

$$\tilde{\omega} = \frac{\tilde{k}_z \tilde{v}_A}{\sqrt{1 + \tilde{k}_y^2 \tilde{\lambda}_e^2}}, \quad (3.76)$$

which is the normalised version of equation (3.34). For a system in which  $\tilde{\lambda}_e = 0.1$  and  $\tilde{v}_A = 1$  are uniform,  $\tilde{k}_y = 8\pi$  and  $\tilde{k}_z = 2\pi$ , (3.76) returns an angular frequency  $\tilde{\omega} = 2.323$ , which

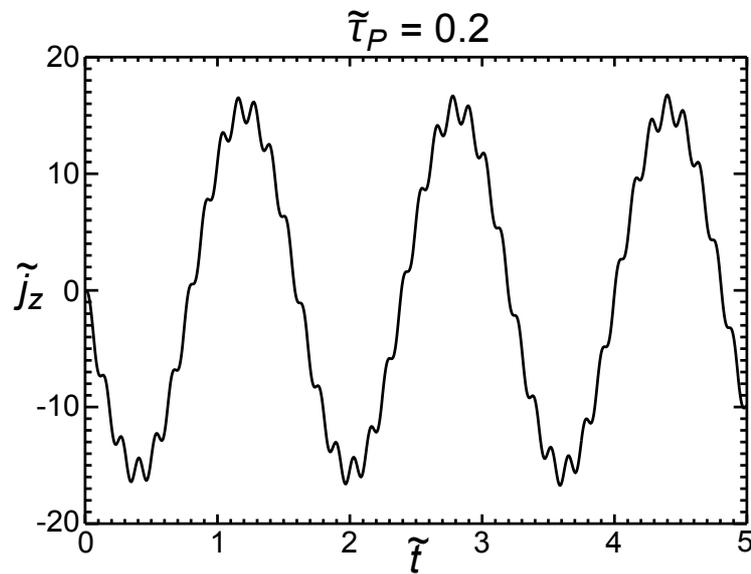


Figure 3.5: Field-aligned current,  $\tilde{j}_z$  at a fixed location for a normal-mode solution with  $\tilde{\tau}_p = 0.2$  everywhere. The long period oscillation is the code's representation of the MHD mode, and the short period oscillation is the artificial plasma oscillation.

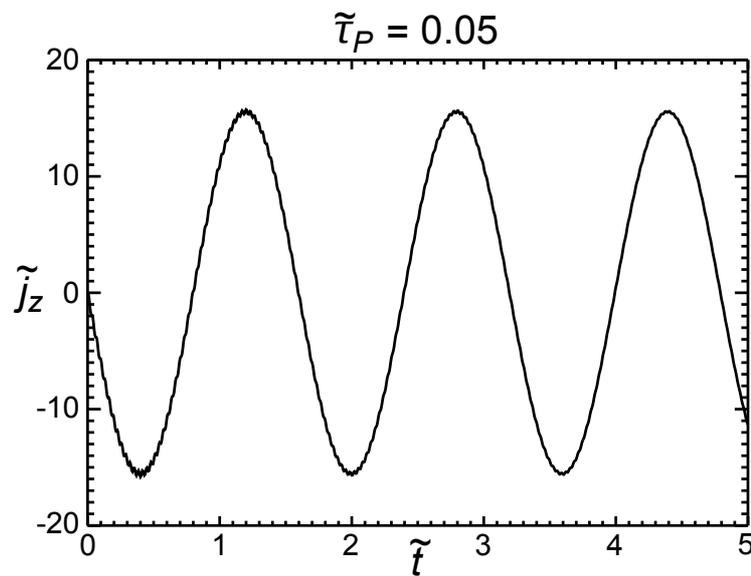


Figure 3.6: Field-aligned current,  $\tilde{j}_z$  at a fixed location for a normal-mode solution with  $\tilde{\tau}_p = 0.05$  everywhere. The short period artificial plasma oscillation has little impact on the long-period MHD solution that dominates  $\tilde{j}_z$ .

corresponds to a period  $\tilde{T} = 2.705$ . Figure 3.7 plots  $\tilde{E}_y$ , as computed by the numerical code for these parameters. Five periods last  $13.525 \pm 0.025$  Alfvén times, corresponding to  $\tilde{T} = 2.705 \pm 0.005$ , in perfect agreement with the dispersion relation.

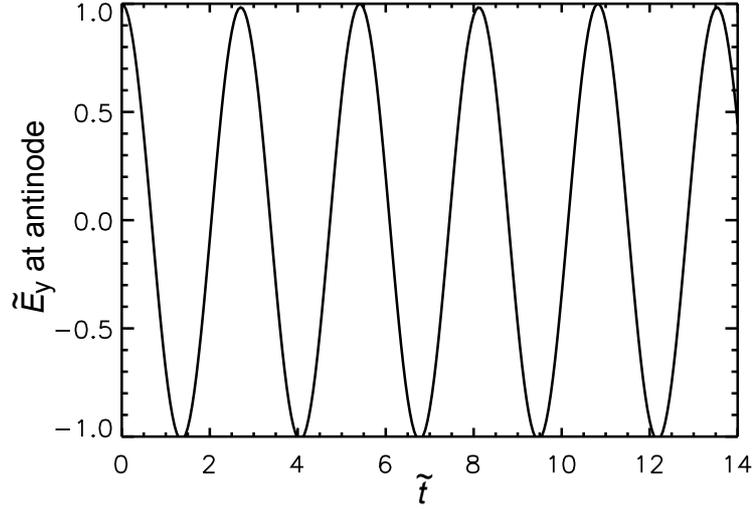


Figure 3.7: Verification of dispersion relation. This plot shows the oscillation in  $\tilde{E}_y$  at a fixed point, for a normal-mode. The dispersion relation predicts a period of  $\tilde{T} = 2.705$ , in excellent agreement with the simulation.

The final test we conducted was conservation of energy for a normal-mode solution. In a cold magnetosphere, energy travels as Poynting flux, so the energy invariant,  $\tilde{\gamma}$ , satisfies

$$\frac{\partial}{\partial \tilde{t}} \iiint_{\text{vol.}} \tilde{\gamma} d\tilde{V} = \iint_{\text{surf.}} \tilde{\mathbf{S}} \cdot \hat{\mathbf{n}} d\tilde{s}, \quad (3.77)$$

where  $\tilde{\mathbf{S}} = \tilde{\mathbf{E}} \times \tilde{\mathbf{B}}$  is the normalised Poynting vector, and  $\hat{\mathbf{n}}$  is a unit vector that is everywhere perpendicular to a closed surface. The double integral on the right-hand side is evaluated over a closed surface, and the triple integral on the left-hand side is evaluated over the volume contained within that surface. This equation states that changes in energy are solely due to net Poynting flux across a boundary. For the normal-mode solution, boundary-conditions prevent Poynting flux from entering or leaving the simulation domain, so total energy within the domain should remain constant.

Using the numerical governing equations (3.68) to (3.71), it can be verified that

$$\tilde{\gamma} = \frac{1}{2} \left( \tilde{b}_x^2 + \frac{\tilde{E}_y^2}{\tilde{v}_A^2} + \frac{\tilde{E}_z^2}{\tilde{c}_{\parallel}^2} + \tilde{\lambda}_e^2 \tilde{J}_z^2 \right) \quad (3.78)$$

satisfies equation (3.77), and we take it as our energy-density. For the run used to test agreement

with the dispersion relation, the integral of energy-density over the domain,  $\tilde{\gamma}_{int}$ , gives

$$\frac{\max(\tilde{\gamma}_{int}) - \min(\tilde{\gamma}_{int})}{\text{mean}(\tilde{\gamma}_{int})} = 2.8 \times 10^{-8}.$$

The difference between  $\max(\tilde{\gamma}_{int})$  and  $\min(\tilde{\gamma}_{int})$  is due to numerical dissipation, a well known artefact that causes total energy to gradually decrease over time when equations are solved on a finite mesh. We therefore conclude that the code conserves energy without any unexpected problems.

### 3.5.6 Incident Inertial Alfvén Wave as an Initial Condition

We aim to place an ionospheric boundary at the bottom of the magnetospheric numerical domain, and drive it by sending in an incident inertial Alfvén wave. An obvious way to drive this wave is to specify wave fields at the top boundary, and various techniques are available to do this. It is better, however, to specify an initial condition within the domain that corresponds to a downgoing wave. The advantages of such an approach are twofold, and arise from the need to prevent reflections from the top boundary interfering with the results.

In order to drive the bottom (ionospheric) boundary for a substantial time, without interference from reflections at the top boundary, the numerical domain is extended in  $\tilde{z}$  compared with its extent in  $\tilde{y}$ ; for example, some of the results in this thesis used an aspect ratio of 25:1. The fastest that a component of an inertial Alfvén wave can travel along the magnetic field is the Alfvén speed,  $\tilde{v}_A$ . Hence, the simulation is halted before a disturbance travelling at this speed can reflect from the top boundary and enter the ‘pure’ part of the domain on which we base any conclusions.

Figure 3.8 illustrates the consideration for the two different types of driving. The left-hand panel of Figure 3.8 shows the incident wave driven at the top boundary. More than a third of the run time is wasted while the wave travels to the bottom boundary from the top. On arrival at  $\tilde{z} = 0$ , the wave partially reflects from the ionosphere, then may later be partially reflected from the top boundary; the simulation must stop before the twice-reflected wave re-enters the ‘pure’ region. Discarding the first third of the data (during which the ionospheric boundary has nothing to interact with) is simply wasteful.

The long time taken for the incident inertial Alfvén wave to reach the bottom boundary produces a second problem too. A large scale disturbance is weakly inertial, so different components of a ramping front travel at slightly different speeds. If the ramp travels over a small distance, then it remains reasonably coherent. However, the large distance separating top and bottom boundaries produces noticeable separation of the ramping front’s constituent parts, and the result is a piecemeal arrival of the incident wave at the ionosphere.

The right-hand panel of Figure 3.8 shows a better alternative: choose an initial condition that produces an incident inertial Alfvén wave, but positions the ramping front near the ionospheric boundary. This reduces the simulation runtime, and produces an Alfvén wave front that arrives at the ionosphere more cleanly.

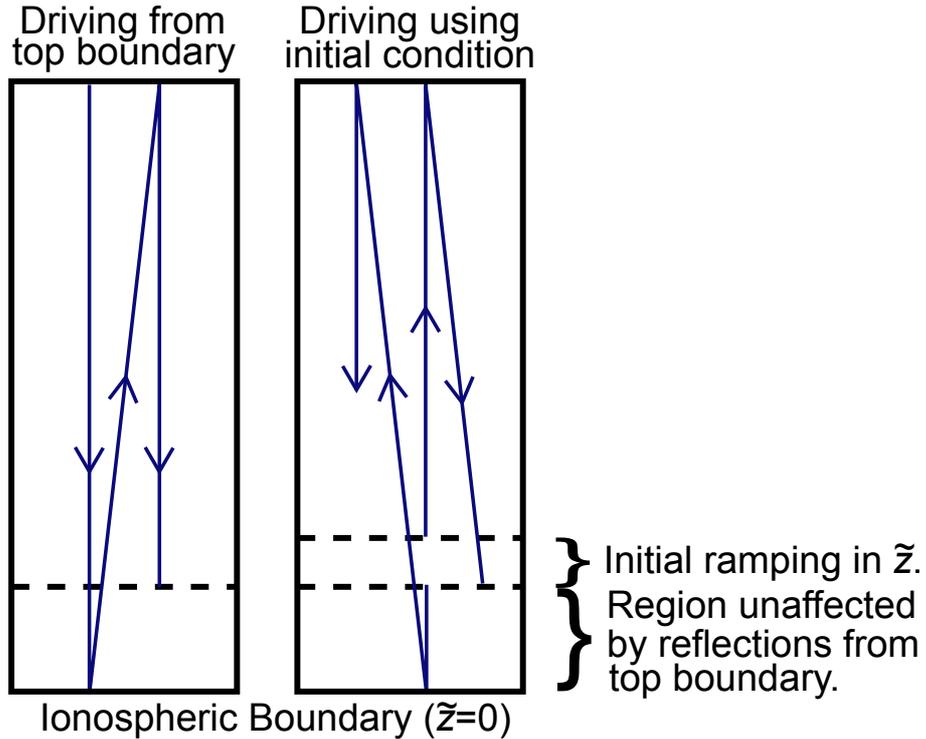


Figure 3.8: Left: An incident Alfvén wave can be sent into the numerical domain by driving the top boundary; the time spent waiting for this wave to reach the bottom ‘ionospheric’ boundary is essentially wasted. Right: A more efficient alternative is to seek an initial condition that approximates an incident IAW, reducing the lost runtime. In both panels, arrows illustrate the maximum distance in  $z$  that waves can travel during a simulation.

The dispersive nature of inertial Alfvén waves makes it extremely difficult (if not impossible) to create a perfect incident inertial-Alfvén-wave with arbitrary variation in  $\tilde{y}$ . It is, however, quite possible to find an approximate solution that serves our purposes admirably well, provided we restrict ourselves to large transverse-scales so that inertial effects are weak. In the limit of weak inertial effects, the ramping front of the incident Alfvén wave contains a weak  $\tilde{E}_z$ , but this does not have a significant effect on the dynamics of the Alfvén wave.

Since the governing equations are solved in a limit where displacement current is small, we neglect the left-hand side of (3.70) and set

$$\tilde{j}_z = -\frac{\partial \tilde{b}_x}{\partial \tilde{y}}. \quad (3.79)$$

In equation (3.68) we simplify for weakly inertial waves by neglecting  $\partial\tilde{E}_z/\partial\tilde{y}$  to get

$$\frac{\partial\tilde{b}_x}{\partial\tilde{t}} = \frac{\partial\tilde{E}_y}{\partial\tilde{z}}. \quad (3.80)$$

Because electron inertia plays a vanishing role in the dynamics of the incident Alfvén wave, equations (3.69) and (3.80) are just those of an ideal Alfvén wave; the incident Alfvén wave will therefore travel at the Alfvén speed, allowing us to replace time derivatives using

$$\frac{\partial}{\partial\tilde{t}} \equiv -\tilde{v}_A \frac{\partial}{\partial\tilde{z}}. \quad (3.81)$$

Thus, (3.69) and (3.80) both give,

$$\tilde{E}_y = -\tilde{v}_A \tilde{b}_x. \quad (3.82)$$

Finally, equation (3.71) becomes

$$\tilde{E}_z = -\tilde{\lambda}_e^2 \tilde{v}_A \frac{\partial\tilde{j}_z}{\partial\tilde{z}}. \quad (3.83)$$

If either  $\tilde{b}_x$  or  $\tilde{E}_y$  is specified for the incident wave at  $\tilde{t} = 0$ , then the above equations determine all of the initial wave fields  $(\tilde{b}_x, \tilde{E}_y, \tilde{E}_z, \tilde{j}_z)$ . The solution can be thought of as a leading order approximation to the true inertial Alfvén wave, and is accurate when length-scales in  $\tilde{y}$  are much larger than  $\tilde{\lambda}_e$ .

The solution is exact where  $\partial/\partial\tilde{z} = 0$ , but only approximate where the incident Alfvén wave ramps in  $\tilde{z}$ . Tests show that the approximate solution works very well, with only one small complaint: the approximate nature of the result produces an additional low-amplitude structure in  $\tilde{E}_z$ , which has an amplitude of a few percent of the largest  $\tilde{E}_z$  in the ramping front, and does not propagate away from the initial location of the ramping front. If the ramp is initially positioned just outside the ‘pure’ region, then this initial condition produces a very clean incident wave, with all of the benefits discussed above.

### 3.5.7 Ionospheric Boundary-Condition

We have established how to solve the magnetospheric governing equations in a physically realistic limit, considered driving, and tested these parts of the model; all of which brings us to the final piece of development: implementation of the ionospheric boundary. There are two parts to this: first, we must keep track of height-integrated number density and alter it in response to field-aligned currents at the top of the E-region; and second, we must update magnetospheric fields as though there were an E-region at the bottom boundary of the numerical domain.

Evolution of height-integrated number density is governed by equation (3.2). Putting

$$\begin{aligned}\tilde{N} &= \frac{N}{N_e}, \\ \eta &= \frac{2B_0\Sigma_{P0}}{eN_e},\end{aligned}$$

and normalising all other quantities as for the magnetospheric equations (Section 3.5.2),  $\tilde{N}$  is governed by

$$\frac{\partial\tilde{N}}{\partial\tilde{t}} = \frac{\epsilon\eta}{2\beta}\tilde{j}_z + \tilde{\alpha}(1 - \tilde{N}^2), \quad (3.84)$$

where

$$\tilde{\alpha} = \frac{\alpha N_e l_0}{h v_{A0}}. \quad (3.85)$$

The constants  $\epsilon\eta$ ,  $\beta$  and  $\tilde{\alpha}$  are all set by the user. Thus, if  $\tilde{j}_z$  is known at the top of the E-region, updating  $\tilde{N}$  is straightforward and can be done using the numerical scheme used for magnetospheric fields.

I will emphasise that the velocities here are normalised with respect to the maximum Alfvén speed and the magnitude of perturbations comes through the parameter  $\epsilon$  in  $b_x = \epsilon B_0 \tilde{b}_x$ . This is in contrast to the 1D code used to model coupling with an ideal magnetosphere, where velocities were normalised with respect to the velocity-amplitude of the incident Alfvén wave. In each code, we use the natural normalisation, for best numerical convenience, but care is needed to compare time-scales or the normalised values of  $\alpha$ .

The second part of ionosphere-magnetosphere interaction is the response of the magnetosphere to the E-region. Wave fields at  $\tilde{z} = 0$  should satisfy the normalised version of equation (3.6), which is

$$\tilde{b}_x = \beta\tilde{N}\tilde{E}_y. \quad (3.86)$$

It is easier to apply this condition if we stagger the grid in  $\tilde{z}$  to separate  $\tilde{b}_x$  and  $\tilde{E}_y$ , as permitted by the pattern of spatial derivatives in the governing equations. The resulting grid is described in Section 3.5.3.2 and illustrated in Figure 3.3.

Staggering the grid so that  $\tilde{b}_x$ ,  $\tilde{E}_z$  and  $\tilde{j}_z$  are solved at  $\tilde{z} = 0$ , and  $\tilde{E}_y$  at  $\tilde{z} = \pm\Delta\tilde{z}/2$ , we interpolate for  $\tilde{E}_y$  at  $\tilde{z} = 0$  with

$$\tilde{E}_y(0) \equiv \frac{\tilde{E}_y(-\Delta\tilde{z}/2) + \tilde{E}_y(+\Delta\tilde{z}/2)}{2}. \quad (3.87)$$

Combining this with equation (3.86), we set  $\tilde{E}_y$  at the ghost-cell as

$$\tilde{E}_y(-\Delta\tilde{z}/2) = \frac{2}{\beta\tilde{N}}\tilde{b}_x(0) - \tilde{E}_y(\Delta\tilde{z}/2). \quad (3.88)$$

Division by  $\beta\tilde{N}$  in (3.88) will lead to numerical instability at the boundary if  $\beta\tilde{N}$  becomes small. Experiments have shown that this instability occurs if  $\beta\tilde{N} \lesssim 0.2$ .

It would have been equally possible to stagger the grid so that  $\tilde{b}_x$  was computed at the ghost-cell using

$$\tilde{b}_x(-\Delta\tilde{z}/2) = 2\beta\tilde{N}\tilde{E}_y(0) - \tilde{b}_x(\Delta\tilde{z}/2), \quad (3.89)$$

but computing  $\tilde{E}_y$  at the ghost-cell is favourable for two reasons. First, if  $\tilde{j}_z$  is defined at  $\tilde{z} = 0$ , then it is directly available to update  $\tilde{N}$  in equation (3.84). Second, if  $\tilde{E}_y$  is computed for the ghost-cell, the boundary remains stable for  $\beta\tilde{N} \gtrsim 0.2$ , a regime that generally suits our purposes. (It is usual to take  $\tilde{N} \approx 1$  at  $\tilde{t} = 0$  and set  $\beta \gtrsim 10$ , giving  $\beta\tilde{N} \gtrsim 10$ .) If  $\tilde{b}_x$  is computed for the ghost-cell, then the boundary is only stable for  $\beta\tilde{N} \lesssim 3$ , which would be less desirable.

Tests were performed to verify proper working of the ionospheric boundary-condition. The first check was reflection of ideal Alfvén waves (independent of  $\tilde{y}$ ) from a passive ionosphere ( $\tilde{N} = 1$  enforced at all times). As discussed in Section 3.3.3, we expect reflection of an ideal Alfvén wave to be governed by

$$\frac{\tilde{E}_r}{\tilde{E}_i} = -\frac{\tilde{b}_r}{\tilde{b}_i} = r = \frac{1 - \beta\tilde{N}}{1 + \beta\tilde{N}}. \quad (3.90)$$

This relationship was clearly seen in tests that covered a large range of  $\beta\tilde{N}$ .

The response of the active ionosphere to field-aligned current was also tested, by ‘switching off’ electron inertia in the magnetosphere and comparing results with similar runs of the 1D code. (An option to neglect inertial effects by updating  $\tilde{b}_x$  and  $\tilde{E}_y$  only, under the assumption  $\tilde{E}_z = 0$ , was built into the code.) Agreement between the two codes was excellent, indicating the successful working of both.

## 3.6 Discussion

The successful testing of both numerical codes concludes this chapter. We are ready to exploit the models, governing equations and computer codes, and shall make frequent use of them in the remaining chapters of this thesis. As a closing comment, let me emphasise the usefulness of complementary approaches and methods. We have developed two models, one using an ideal

magnetosphere, and the other including electron inertial effects in the magnetosphere. Both shall show their individual value in the coming chapters, but it is arguably the ability to compare and contrast these models that has produced the most original insights into magnetosphere-ionosphere coupling.

# Self-Consistent Steady-States of the Ionosphere-Magnetosphere System

---

## 4.1 Introduction

Steady-states contain a lot of information about the coupled ionosphere-magnetosphere system, including vital clues to its dynamic behaviour. This chapter takes a detailed look at steady-states produced with an ideal magnetosphere, focusing on the steady-state limit of equation (3.46):

$$\frac{1}{\mu_0 e} \frac{d}{dy} \left( \frac{2\beta N b_i}{N_e + \beta N} \right) = \frac{\alpha}{h} (N_e^2 - N^2). \quad (4.1)$$

We will follow two complementary lines of investigation. First, the computer code described in Section 3.4 is used to obtain sample steady-states; these provide an intuitive feel for steady-states, and also allow simple parametric studies. These results assist with the development of an analytic treatment, in which steady-state solutions are constructed mathematically. Finally, we confirm that numerical and analytic solutions are in excellent agreement.

## 4.2 Computer Simulations

We begin by investigating self-consistent magnetosphere-ionosphere steady-states, obtained using the computer code described in Section 3.4.

### 4.2.1 Setup

The simulated E-region is initialised with uniform height-integrated number density  $\tilde{N} = 1$ , and driven by an incident Alfvén wave with velocity perturbation

$$\tilde{u}_i(\tilde{y}) = \begin{cases} -(1 + \cos(\tilde{y})) & \text{for } -\pi < \tilde{y} < \pi, \\ 0 & \text{otherwise,} \end{cases} \quad (4.2)$$

that shears the background magnetic field to produce field-aligned current. This particular profile produces a channel of upward field-aligned current between  $\tilde{y} = -\pi$  and  $\tilde{y} = 0$ , and an adjacent channel of downward field-aligned current between  $\tilde{y} = 0$  and  $\tilde{y} = \pi$ . The system fits nicely into a numerical domain extending from  $\tilde{y} = -4$  to  $\tilde{y} = 4$ , and because  $d\tilde{u}_i/d\tilde{y} = 0$  gives zero current at the boundaries,  $\tilde{N} = 1$  is a convenient boundary-condition in  $\tilde{y}$ .

Three parameters control the simulation —  $\eta$ ,  $\beta$  and  $\tilde{\alpha}$ . Taking  $B_0 = 5 \times 10^4$  nT,  $\Sigma_{P0} = 1.95$  mho and  $N_e = 1.2 \times 10^{15} \text{ m}^{-2}$  as typical values, equation (3.54) gives  $\eta = 1.015$ , which is used for all of the ideal simulations presented in this chapter and Chapter 6. The ratio,  $\beta$ , of initial height-integrated Pedersen conductance to the Alfvén conductance, is less well defined, so it is varied across a number of simulations. Some time will be spent looking at how the final steady-state depends on  $\beta$ , exploring the range  $\beta = 10$  to  $\beta = 1370$ .

The third parameter,  $\tilde{\alpha}$ , is given by equation (3.55) as

$$\tilde{\alpha} = \frac{\alpha N_e l_0}{h u_{i0}},$$

which contains a mixture of physical values and normalisation quantities on the right-hand side. The physical values are easily set, and we use  $\alpha = 3 \times 10^{-13} \text{ m}^3 \text{ s}^{-1}$  and  $h = 2 \times 10^4$  m. This leaves the ratio  $\tau = l_0/u_{i0}$  to be determined. If we define

$$j_{||0} = \left| \min \left( -\frac{1}{\mu_0} \frac{d}{dy} b_x(y, t = 0) \right) \right| \quad (4.3)$$

as the unsigned extreme current-density in the downward current-channel at  $t = 0$ , then we can use equations (3.43) and (3.44), together with  $N(y, t = 0) = N_e$ , to write

$$\begin{aligned} j_{||0} &= \frac{1}{\mu_0} \max \left( \frac{d}{dy} ((1 - r_0) b_i) \right) \\ \Rightarrow j_{||0} &= \frac{1}{\mu_0 l_0} \max \left( \frac{d}{d\tilde{y}} \left( \frac{2\beta}{(1 + \beta)} \frac{B_0 u_{i0} \tilde{u}_i}{v_A} \right) \right) \\ \Rightarrow j_{||0} &= \frac{B_0 u_{i0} \beta}{\mu_0 v_A l_0 (1 + \beta)} \max \left( \frac{d\tilde{u}_i}{d\tilde{y}} \right). \end{aligned}$$

Since  $\max(d\tilde{u}_i/d\tilde{y}) = 1$  for our chosen driver, we can rearrange the above to get

$$\Rightarrow \tau = \frac{l_0}{u_{i0}} = \frac{2B_0 \Sigma_{P0}}{(1 + \beta) j_{||0}}, \quad (4.4)$$

which determines  $\tilde{\alpha}$  as

$$\tilde{\alpha} = \frac{2\alpha N_e B_0 \Sigma_{P0}}{h(1 + \beta) j_{||0}}.$$

Unless otherwise stated, simulations use  $j_{||0} = 5 \mu\text{Am}^{-2}$ .

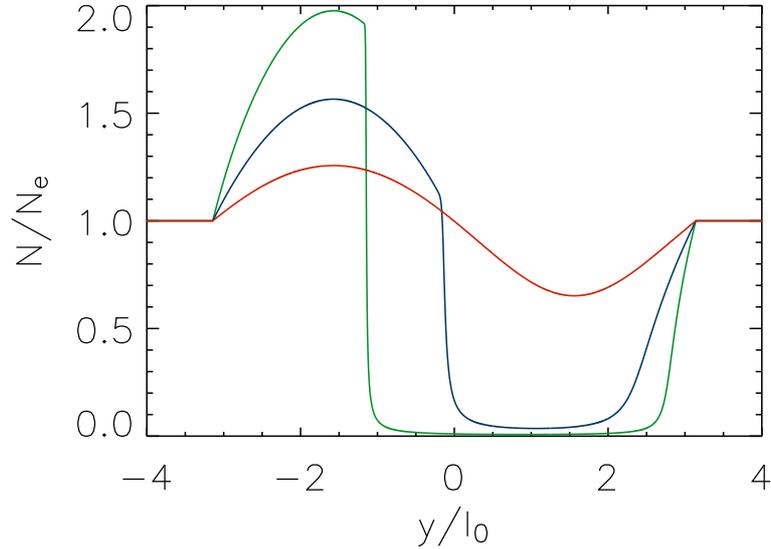


Figure 4.1: Numerical steady-states for three different initial current densities:  $j_{||0} = 2 \mu\text{Am}^{-2}$  (red),  $5 \mu\text{Am}^{-2}$  (blue) and  $10 \mu\text{Am}^{-2}$  (green). The region of suppressed E-region plasma-density is deeper and wider for increased initial current density.

This chapter focuses exclusively on the steady-states toward which simulations converge, postponing discussion of dynamics (for the most part) until Chapter 6. Here convergence is obtained by running all simulations until  $\left| \int (1 - \tilde{N}^2) d\tilde{y} \right|$  is less than  $1.0 \times 10^{-10}$  and slowly converging to zero, as evaluated over the simulation domain (see Section 3.4.4 for more information about this condition).

#### 4.2.2 Overview of Steady-States

Figure 4.1 shows steady-states for  $\beta = 100$  and three different values of  $j_{||0}$ :  $2 \mu\text{Am}^{-2}$ ,  $5 \mu\text{Am}^{-2}$  and  $10 \mu\text{Am}^{-2}$ . The coarse features are just as one would expect: in the upward current-channel, electrons are deposited in the E-region, increasing  $\tilde{N}$ , whereas in the downward current-channel, electrons are removed from the E-region, decreasing  $\tilde{N}$ . However, if  $j_{||0}$  is sufficiently large (as in the cases  $j_{||0} = 5 \mu\text{Am}^{-2}$  and  $10 \mu\text{Am}^{-2}$ ), ionospheric depletion becomes so severe that the downward current-channel is forced to broaden into the upward current-channel.

If  $j_{||0}$  is now held constant (at  $j_{||0} = 5 \mu\text{Am}^{-2}$ ), we can study how the solution changes with different values of  $\beta$ . Three steady-states, obtained for  $\beta = 20, 100$  and  $1370$ , are shown in Figure 4.2. On the whole, the steady-states are very similar, but two differences are readily apparent: first, as  $\beta$  increases, the minimum density decreases; and second, greater values of  $\beta$  shorten the

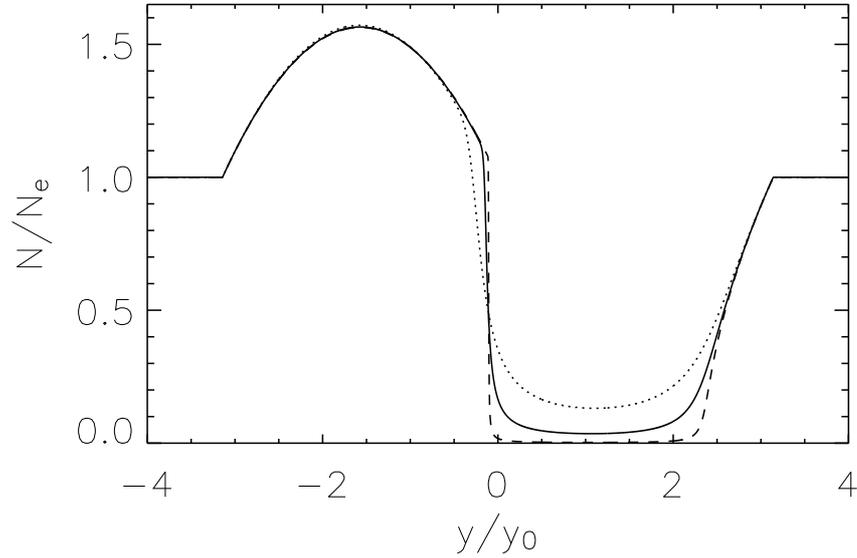


Figure 4.2: Numerical steady-states for  $\beta = 20$  (dotted curve),  $\beta = 100$  (solid curve) and  $\beta = 1370$  (dashed curve). As  $\beta$  is increased, the suppression of number densities in the downward current-channel becomes more severe. The shortest length-scale in each steady-state decreases with increasing  $\beta$ , leading to steeper gradients. Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

smallest length-scale, which is located at the steep transition in  $\tilde{N}$  in between the two current-channels.

### 4.2.3 Minimum E-Region Plasma-Density

It is clear from Figure 4.2 that the minimum number density,  $N_{min}$ , decreases with  $\beta$ , but the precise form of the relationship is unclear from this figure alone. The relationship does become clear though, if we examine a larger number of examples. Figure 4.3 shows data for 12 different values of  $\beta$ , spanning the range  $\beta = 10$  to 400. The inverse of  $\tilde{N}_{min}$  is clearly linear in  $\beta$  over most of this range, and the best fit straight line is given by  $2.64 + 0.253\beta$ . It follows that

$$\tilde{N}_{min} \approx \frac{1}{2.64 + 0.253\beta} \approx \frac{3.95}{\beta(1 + 10.4/\beta)} \quad (4.5)$$

for these simulations. For large  $\beta$ , the bracket goes to unity, so the main dependence is  $\tilde{N}_{min}$  proportional to  $1/\beta$ . This dependence is revisited in Section 4.3.4 using analytic tools.

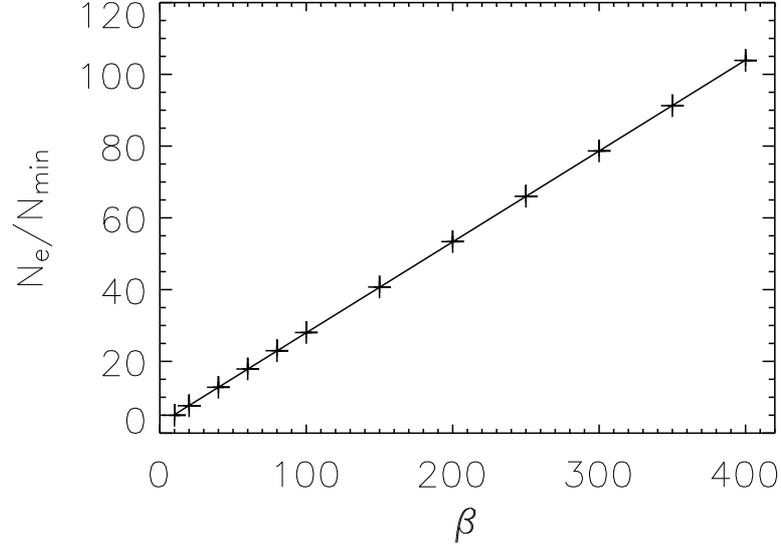


Figure 4.3: Demonstration that  $N_e/N_{min}$  is linear in  $\beta$  for the numerical steady-states. Crosses represent data from simulations and the straight line is a best fit to the data, given by  $2.64+0.253\beta$ . Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

#### 4.2.4 Shortest Length-Scale in Steady-state

The same set of studies can be used to examine the shortest length-scale in the steady-state, which is inversely proportional to the steepest gradient. Figure 4.4 shows the dependence on  $\beta$ : the gradient is linear in  $\beta$  over most of the range  $\beta = 10$  to 400, and the best fit straight line is given by  $-1.42 - 0.108\beta$ . If the change in  $\tilde{N}$  between the upward and downward current channels is  $\Delta\tilde{N} \approx 1$ , then we can estimate the shortest length-scale from

$$\tilde{l}_{min} \approx \frac{\Delta\tilde{N}}{\left|d\tilde{N}/d\tilde{y}\right|_{\max}} \approx \frac{1}{1.42 + 0.108\beta} \approx \frac{9.26}{\beta(1 + 13.1/\beta)}, \quad (4.6)$$

which is proportional to  $1/\beta$  for large  $\beta$ . This feature is revisited in Section 4.4.9 using analytic tools.

### 4.3 Analytic Solutions

The steady-state equation (4.1) is readily solved in two limits. The first of these assumes that  $\beta N \gg N_e$ , and is valid where depletion has not significantly altered the reflection coefficient from

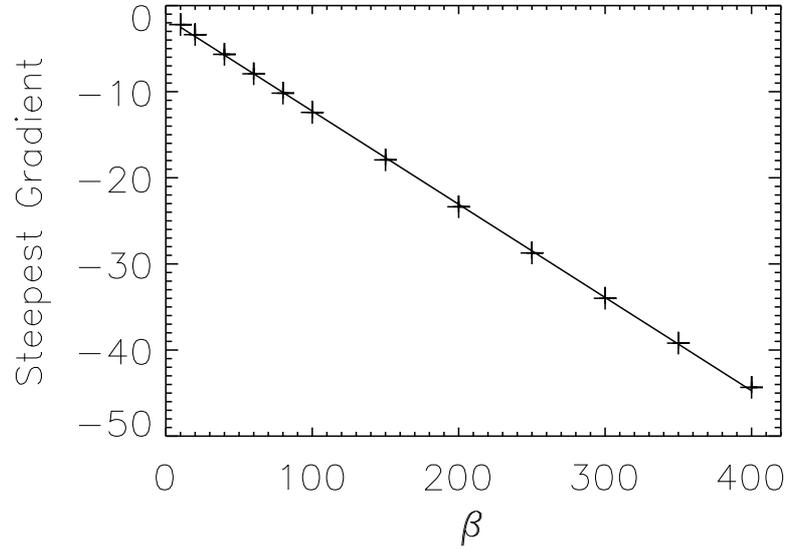


Figure 4.4: Demonstration that the steepest gradient in the numerical steady-state is linear in  $\beta$ . Crosses represent data from simulations and the straight line is a best fit to the data, given by  $-1.42 - 0.108\beta$ . Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

$r = -1$  (the ionosphere remains highly reflective). This leads to an upper steady-state solution,  $N_{upper}$ . The second limit assumes that  $N^2 \ll N_e^2$ ; this corresponds to significant depletion, so that recombination is negligible, and leads to a lower solution,  $N_{lower}$ . Once upper and lower solutions have been obtained, the global solution can be constructed through boundary-layer matching.

### 4.3.1 Upper Solution

If  $\beta N \gg N_e$ , equation (4.1) reduces to

$$\frac{2}{\mu_0 e} \frac{db_i}{dy} = \frac{\alpha}{h} (N_e^2 - N^2), \quad (4.7)$$

which is easily rearranged for the upper steady-state,

$$N_{upper} = N_e \sqrt{1 - \frac{2}{\alpha n_e^2 h e \mu_0} \frac{db_i}{dy}}. \quad (4.8)$$

The result is tidier and makes better physical sense if rewritten in terms of field-aligned current densities. Using Ampère's law, the vertical current density in the incident wave is

$$j_i = -\frac{1}{\mu_0} \frac{db_i}{dy}. \quad (4.9)$$

Meanwhile, ionisation in the E-region produces  $\alpha n_e^2 h$  electrons per unit area and time, equivalent to an available 'ionisation current density',

$$j_c = \alpha n_e^2 h e. \quad (4.10)$$

Hence, the upper steady-state can be expressed as

$$N_{upper} = N_e \sqrt{1 + \frac{2j_i}{j_c}}. \quad (4.11)$$

The switch between real  $N_{upper}$  ( $2j_i/j_c > -1$ ) and imaginary  $N_{upper}$  ( $2j_i/j_c < -1$ ) is physically significant, and is explained as follows. We have assumed  $\beta N \gg N_e$ , so that the reflection coefficient is not significantly altered from  $r = -1$ ; therefore (since  $r$  is uniform) the total field-aligned current density is  $j_z = (1 - r)j_i = 2j_i$ . In downward current regions, electrons are drawn from the E-region, which can supply a maximum current density of  $j_c$ : if a downward current draws more than the ionisation current density (i.e. if  $2j_i < -j_c$ ) then electrons are removed from the ionosphere until the reflection coefficient *does* change significantly from  $r = -1$ , and the total current density is reduced to a sustainable level. At any such locations the upper solution inevitably breaks down, and a lower solution is required.

The above discussion leads us to consider broadening of the downward current-channel. In the upward current-channel, E-region number density is enhanced, so the reflection coefficient remains  $r = -1$ , and this is uniform except possibly at the border with the downward current-channel. Hence, the total field-aligned current remains constant in most of the upward current-channel (the magnetosphere acts as an essentially unlimited supply of electrons, so upward current is easily sustained). If  $|2j_i| > j_c$  within the downward current-channel, then a reflection coefficient of  $r = -1$  cannot be sustained: here, E-region number density becomes strongly suppressed, reducing the total field-aligned current density. But, the total (integrated) current over the upward current region must be balanced by the total downward current, so that currents close and the E-region remains quasi-neutral. The downward current region is therefore forced to expand, increasing the area of ionosphere from which it can draw the reduced current density, until total upward and downward currents match. In this way, the criterion for failure of the upper solution ( $|2j_i| > j_c$  at some point in a downward current region) also tests for broadening of the downward current-channel.

The change of behaviour that occurs at  $|2j_i| = j_c$  for downward current was previously noted by Cran-McGreehin et al. (2007), who introduced  $j_c$  as a critical current density for broadening. Their work was based on a more restricted set of steady-states than we consider here, so the present work generalises this important result. (This usage of  $j_c$  (as a critical current density that we compare to twice the incident current density,  $2j_i$ ) is consistent with the original definition introduced by Cran-McGreehin et al. (2007); it is not, however, the same as the  $j_c$  that appears in Russell et al. (2010), which was defined for comparison to the total (incident plus reflected) current-density at  $t = 0$ .)

### 4.3.2 Lower Solution

Where  $N^2 \ll N_e^2$ , equation (4.1) reduces to

$$\frac{d}{dy} \left( \frac{\beta N b_i}{N_e + \beta N} \right) = \frac{\mu_0 e \alpha N_e^2}{2h} = \frac{\mu_0 j_c}{2}. \quad (4.12)$$

Integrating this directly and rearranging yields the lower steady-state solution

$$N_{lower} = \frac{N_e(y + c)}{\beta(\Omega b_i - (y + c))}, \quad (4.13)$$

where  $c$  is an integration constant and

$$\Omega = \frac{2}{\mu_0 j_c} = \frac{2}{\mu_0 \alpha n_e^2 h e} \quad (4.14)$$

is also constant.

The lower solution breaks down at points where the denominator goes to zero, i.e. where

$$c = \Omega b_i(y) - y. \quad (4.15)$$

Therefore,  $c$  can be determined if there is a known location at which the lower solution must break down. The dynamics of the coupled magnetosphere-ionosphere system allow the downward current-channel to expand in the direction of the incident electric field but not against it (see Chapter 6). If we assume that Pedersen currents close field-aligned currents locally, and consider an upward-downward pair of field-aligned currents, then the downward current region broadens on the side adjacent to the upward current region. This allows us to fix the value of  $c$ , because a global steady-state that has broadened on one side can be constructed if, and only if, upper and lower solutions break down at a common point, on the side of the density trough that has not broadened. From the upper steady-state, this location is the point farthest from the upward current

region at which,

$$j_i(y) = -\frac{j_c}{2}, \quad (4.16)$$

and we obtain  $c$  by substituting this value of  $y$  into equation (4.15).

### 4.3.3 Critical Points and Integration Constant from Plot of $\Omega b_i(y) - y$

A plot of  $\Omega b_i(y) - y$  against  $y$  is a powerful visual tool that quickly finds critical points on the upper and lower solutions and determines the value of integration constant  $c$ . An illustration is given in Figure 4.5. From equation (4.11), the upper steady-state breaks down at critical points where  $j_i(y) = -j_c/2$ . Since

$$\begin{aligned} j_i(y) &= -\frac{j_c}{2} \\ \Rightarrow -\frac{1}{\mu_0} \frac{db_i}{dy} &= \frac{j_c}{2} \\ \Rightarrow \frac{2}{j_c \mu_0} \frac{db_i}{dy} - 1 &= 0 \\ \Rightarrow \frac{d}{dy} (\Omega b_i(y) - y) &= 0, \end{aligned}$$

it follows that these critical points are the turning points of  $\Omega b_i(y) - y$ . Meanwhile, the lower steady-state is singular where

$$\Omega b_i(y) - y = c,$$

i.e. at intersections with a horizontal line whose vertical position is  $c$ . The condition that the downward current-channel broadens on one side requires  $f(y) = c$  to intersect the turning point of  $f(y) = \Omega b_i - y$  farthest from the upward current region (which is always a maximum). Therefore, the value of  $c$  is readily determined, providing a unique lower solution.

### 4.3.4 Minimum E-Region Plasma-Density

If  $|2j_i| < j_c$  in all downward current regions (and  $\beta \gg 1$ ) then the upper solution gives a complete description of the steady-state. In this case, the minimum E-region plasma-density is located at the point of greatest downward current density, and

$$N_{min} = N_e \sqrt{1 + \min \left( \frac{2j_i}{j_c} \right)}. \quad (4.17)$$

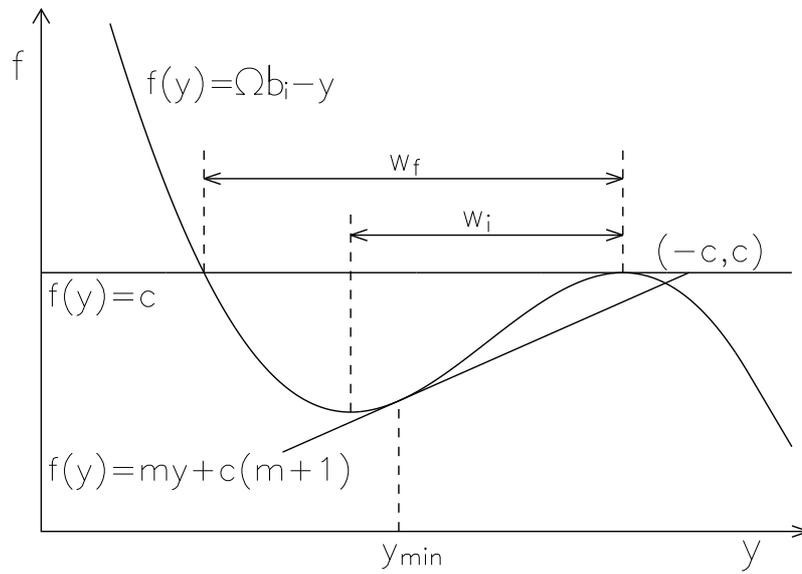


Figure 4.5: Sketch of  $\Omega b_i(y) - y$  against  $y$ , from which many key features of the solution can be determined. The upper solution breaks down at turning points of  $\Omega b_i(y) - y$ . The horizontal line, passing through the maximum turning point, determines the integration constant,  $c$ , and the lower solution breaks down where this horizontal line intersects  $\Omega b_i(y) - y$ . A straight line through the point  $(-c, c)$ , that is tangent to  $\Omega b_i - y$  between its turning points, reveals the minimum number density: the tangent meets the curve at the location of the minimum in number density,  $y_{min}$ ; and the minimum number density is  $N_e/\beta m$  where  $m$  is the slope of the tangent. The horizontal distance between the turning points,  $w_i$ , indicates the initial width of the plasma-density cavity, and the distance between the intersections  $\Omega b_i - y = c$ , indicates the final width ( $w_f$ ) which it attains by broadening. Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

If  $|2j_i| > j_c$  at some point in a downward current region, then the density minimum is obtained from the lower solution. Writing  $y_{min}$  as the location of the density minimum, putting equation (4.13) into  $[dN_{lower}/dy]_{y=y_{min}} = 0$  gives

$$b_i(y_{min}) - b'_i(y_{min})(y_{min} + c) = 0, \quad (4.18)$$

where a prime denotes differentiation with respect to  $y$ . Using this to replace  $b_i$  in (4.13), the minimum E-region plasma-density is

$$N_{min} = \frac{N_e}{\beta (\Omega b'_i(y_{min}) - 1)}. \quad (4.19)$$

The main dependence on  $\beta$  is  $N_{min} \sim 1/\beta$ , in keeping with the numerical results presented in Section 4.2.3.

We have already seen that a plot of  $\Omega b_i - y$  against  $y$  (Figure 4.5) can be used to find the integration constant  $c$ ; the same plot also provides a very direct means of getting the minimum density when  $|2j_i| > j_c$  in a downward current region. At  $y_{min}$ , the tangent to  $\Omega b_i - y$  has gradient

$$m = \frac{d}{dy} (\Omega b_i - y) = \Omega b'_i(y_{min}) - 1. \quad (4.20)$$

Therefore, multiplying (4.18) through by  $\Omega$ , putting  $\Omega b'_i(y_{min}) = m + 1$ , and rearranging for  $m$ , we can write

$$m = \frac{c - (\Omega b_i(y_{min}) - y_{min})}{-c - y_{min}}. \quad (4.21)$$

Because the tangent is a straight line, we can also write

$$m = \frac{f_2 - f_1}{y_2 - y_1}, \quad (4.22)$$

where  $(y_1, f_1)$  and  $(y_2, f_2)$  are two points that lie on the tangent. Therefore, putting  $(y_1, f_1) = (y_{min}, \Omega b_i(y_{min}) - y_{min})$  into (4.22) and comparing with equation (4.21), we see that the tangent to  $\Omega b_i - y$  at  $y_{min}$  passes through  $(-c, c)$ . Conversely, a straight line that is tangent to  $\Omega b_i - y$  between its turning points, and passes through  $(-c, c)$  meets  $\Omega b_i - y$  at  $y_{min}$ . Therefore, adding a single straight line to the plot gives us the location of the density minimum. Furthermore, putting (4.20) into (4.19), we are free to write

$$N_{min} = \frac{N_e}{\beta m}, \quad (4.23)$$

where  $m$  is the slope of the tangent. Both the value and location of the density minimum have now been found using the plot.

### 4.3.5 Plasma-Density Cavity: Existence, Width and Broadening

What does it mean to say that a density cavity has formed? A number of possible definitions are available, but consideration of steady-states has led me to favour the following: *a region of suppressed E-region plasma-density should be considered a cavity if, and only if, the reflection coefficient is significantly altered from  $r = -1$* . Under this definition, existence of an E-region plasma-density cavity is significant for the magnetosphere, putting the definition on a good physical basis for magnetosphere-ionosphere studies. This is much more informative than an arbitrary cut-off in E-region plasma-density. In the context of steady-states, a density cavity exists wherever a lower solution is required. Hence, from the discussion of Section 4.3.1, a cavity forms if  $|2j_i| > j_c$  for downward current, and it broadens as the downward current-channel broadens.

Steady-state solutions can be used to estimate both the initial width of a density cavity, and the final width which it attains by broadening. The initial width is estimated as the distance between the two points at which the upper solution breaks down, i.e. at which  $\Omega b'_i(y) = 1$ . Inspection of a plot of  $\Omega b_i(y) - y$  against  $y$  (Figure 4.5) yields this as the horizontal distance between the two turning points (marked as  $w_i$ ).

Similarly, the final width is estimated as the distance between the two points at which the lower solution is singular, i.e. at which  $\Omega b_i - y = c$ . Referring to the plot of  $\Omega b_i(y) - y$  against  $y$  (Figure 4.5), this final width (marked as  $w_f$ ) is the horizontal distance between intersections of  $f(y) = \Omega b_i(y) - y$  and  $f(y) = c$ . It can be seen that, regardless of the value of  $c$ , the final width increases monotonically with  $\Omega$ . Writing  $y = a$  and  $y = b$  for the limits of the plasma-density cavity, where  $b > a$ , the final width satisfies

$$w_f = b - a = \Omega (b_i(b) - b_i(a)); \quad (4.24)$$

the strong dependence on  $\Omega$  is readily apparent. Recalling that  $b_i \rightarrow 0$  at the edge of the current system (for a localised disturbance that does not allow currents to close at infinity), it is also clear that the plasma-density cavity cannot expand beyond the limits of the current system.

## 4.4 Global Analytic Solution By Boundary-Layer Matching

The global steady-state can be accurately approximated by using a boundary-layer method to join the upper and lower steady-state solutions. The principle is that the upper steady-state is accurate where there is little depletion, and the lower steady-state is accurate where depletion is significant, leaving two boundary-layers in which neither approximation is valid and the solution makes a transition between upper and lower steady-states (see Figure 4.6). Within each boundary-layer, we can make progress with the original steady-state equation (4.1) if we assume that these layers are

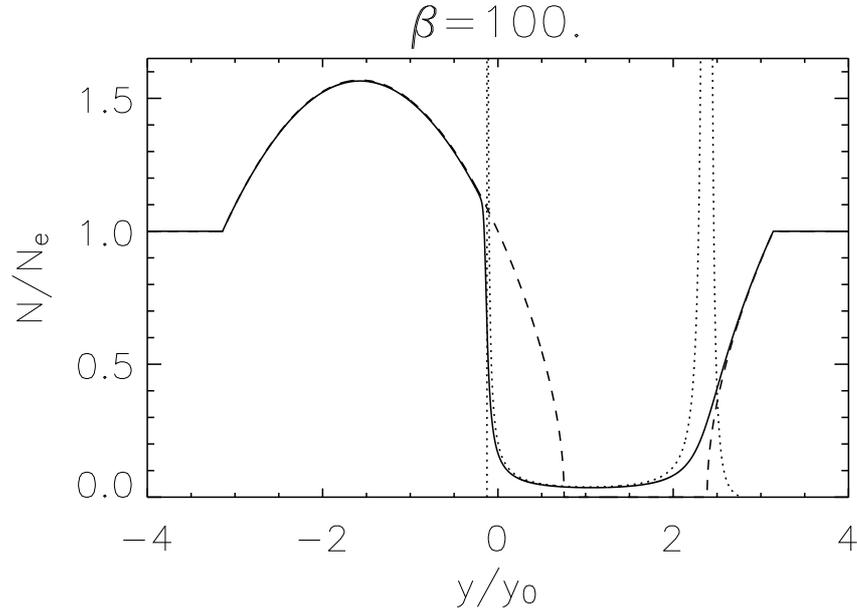


Figure 4.6: Plot showing the steady-state obtained numerically for  $\beta = 100$  (solid line) with the upper steady-state (dashed curve) and lower steady-state (dotted curve). Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

narrow, and that solutions only need to be valid within the boundary-layer; then, having obtained boundary-layer solutions, we can use them to construct the global steady-state.

#### 4.4.1 Overview of Boundary-Layer Method

For a boundary-layer positioned at  $y = \xi$  (a location at which the lower solution breaks down), the boundary-layer matching is performed as follows:

1. Introduce a scaling of the form

$$Y = \beta^\epsilon (y - \xi), \quad (4.25)$$

$$\mathcal{N} = \beta^\nu \frac{N}{N_e}, \quad (4.26)$$

where  $\nu$  may be zero but  $\epsilon > 0$ . This provides a stretched coordinate,  $Y$ , that typically varies from about  $Y = -1$  to  $Y = 1$  across the boundary-layer.

2. Expand  $\mathcal{N}_{upper}^2(Y)$  and  $\mathcal{N}_{lower}(Y)$  about  $\xi$ .

Since  $b_i$  is regular at  $\xi$ , it has a Taylor expansion,

$$b_i(Y) = b_i(\xi) + b'_i(\xi)\beta^{-\epsilon}Y + \frac{b''_i(\xi)}{2}\beta^{-2\epsilon}Y^2 + \dots, \quad (4.27)$$

where a prime denotes differentiation with respect to  $y$ .

$\mathcal{N}_{upper}(Y)$  is readily obtained from  $\mathcal{N}_{upper}^2(Y)$ , and inspection reveals the behaviour of the upper and lower solutions as  $Y \rightarrow 0$ . At this stage, the values of  $\epsilon$  and  $\nu$  can be determined, as described in Section 4.4.2.

### 3. Expand $\mathcal{N}$ .

Led by the occurrence of powers of  $\beta^{-\epsilon}$  in the expansion of  $b_i$ , we expand  $\mathcal{N}$  as

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1\beta^{-\epsilon} + \mathcal{N}_2\beta^{-2\epsilon} + \dots \quad (4.28)$$

If  $\beta^{-\epsilon} < 1$  then  $\mathcal{N}_0$  is an approximation for  $\mathcal{N}$ , while  $\mathcal{N}_1$  etc. provide corrections. The smaller the value of  $\beta^{-\epsilon}$ , the better the approximation.

### 4. Substitute expansions of $\mathcal{N}_{upper}^2$ , $b_i$ and $\mathcal{N}$ into the scaled governing equation, and equate terms in $\beta^0$ .

This gives an ordinary differential equation for  $\mathcal{N}_0$ . Solutions for  $\mathcal{N}_1$  and higher order corrections are obtained by equating lower powers of  $\beta$  in the scaled governing equation.

### 5. Construct the global solution across the boundary-layer.

The global solution is constructed in piecewise fashion, with solutions for  $Y < 0$  and  $Y > 0$  specified separately. For each case ( $Y < 0$  and  $Y > 0$ ) in turn, we set  $\mathcal{N}_{outer}$  as the steady-state solution that is valid outside the boundary-layer (either  $\mathcal{N}_{upper}$  or  $\mathcal{N}_{lower}$ ), and introduce  $g(Y)$  as the leading order behaviour of  $\mathcal{N}_{outer}$  as  $Y \rightarrow 0$ , making sure that  $g(Y)$  includes all singular terms. We also put

$$\mathcal{N}_{inner} = \mathcal{N}_0 + \dots + \mathcal{N}_{r-1}\beta^{-(r-1)\epsilon}, \quad (4.29)$$

where  $r \in \mathbb{N}$  matches or exceeds the number of singular terms in  $\mathcal{N}_{outer}$ .

The inner and outer solutions on one side of  $Y = 0$  can be joined together, provided  $\mathcal{N}_{inner} \rightarrow g(Y)$  as  $|Y| \rightarrow \infty$  in the direction of the outer solution; that is to say, the behaviour of the inner solution as  $|Y| \rightarrow \infty$  towards the outer solution matches the behaviour of the outer solution as  $Y \rightarrow 0$ . The solutions are then joined using

$$\mathcal{N} \approx \mathcal{N}_{inner} + \mathcal{N}_{outer} - g(Y) \quad (4.30)$$

to perform the matching.

6. Rewrite the complete solution in terms of the original, unscaled variables  $N$  and  $y$ .

#### 4.4.2 Scaling of Governing Equation

In order to perform the boundary-layer analysis, we use a stretched coordinate,

$$Y = \beta^\epsilon (y - \xi).$$

It may prove necessary to scale the density as well, so we write

$$\mathcal{N} = \beta^\nu \frac{N}{N_e},$$

noting that  $\nu$  may be zero.

Tidying multiple factors into  $\Omega$  (4.14), the steady-state equation (4.1) may be written as

$$\frac{d}{dy} \left( \frac{\beta N b_i}{N_e + \beta N} \right) = \frac{1}{\Omega} \left( 1 - \frac{N^2}{N_e^2} \right). \quad (4.31)$$

The upper solution satisfies this equation in the limit  $\beta N \gg N_e$ , hence

$$\frac{db_i}{dy} = \frac{1}{\Omega} \left( 1 - \frac{N_{upper}^2}{N_e^2} \right), \quad (4.32)$$

a result that is also clear from (4.11) and the definition of  $\Omega$ . Using the product rule to expand the derivative in (4.31),

$$\frac{d}{dy} \left( \frac{\beta N}{N_e + \beta N} \right) b_i + \left( \frac{\beta N}{N_e + \beta N} \right) \frac{db_i}{dy} = \frac{1}{\Omega} \left( 1 - \frac{N^2}{N_e^2} \right), \quad (4.33)$$

and noting that

$$\frac{\beta N}{N_e + \beta N} = 1 - \frac{1}{1 + \beta N/N_e}, \quad (4.34)$$

further simplification gives

$$-\frac{d}{dy} \left( \frac{1}{1 + \beta N/N_e} \right) b_i + \left( 1 - \frac{1}{1 + \beta N/N_e} \right) \frac{db_i}{dy} = \frac{1}{\Omega} \left( 1 - \frac{N^2}{N_e^2} \right) \quad (4.35)$$

$$\Rightarrow \frac{\beta b_i}{(1 + \beta N/N_e)^2} \frac{d}{dy} \left( \frac{N}{N_e} \right) - \frac{1}{1 + \beta N/N_e} \frac{db_i}{dy} = \frac{1}{\Omega} \left( 1 - \frac{N^2}{N_e^2} \right) - \frac{db_i}{dy} \quad (4.36)$$

$$\Rightarrow \frac{\beta b_i}{(1 + \beta N/N_e)^2} \frac{d}{dy} \left( \frac{N}{N_e} \right) - \frac{1}{1 + \beta N/N_e} \frac{db_i}{dy} = \frac{1}{\Omega} \left( \frac{N_{upper}^2}{N_e^2} - \frac{N^2}{N_e^2} \right) \quad (4.37)$$

$$\Rightarrow \beta b_i \frac{d}{dy} \left( \frac{N}{N_e} \right) - \left( 1 + \beta \frac{N}{N_e} \right) \frac{db_i}{dy} = \frac{1}{\Omega} \left( 1 + \beta \frac{N}{N_e} \right)^2 \left( \frac{N_{upper}^2}{N_e^2} - \frac{N^2}{N_e^2} \right). \quad (4.38)$$

Applying the scaling ( $N/N_e = \beta^{-\nu} \mathcal{N}$  and  $dy \equiv \beta^{-\epsilon} dY$ ), this becomes,

$$\begin{aligned} \beta^{1+\epsilon-\nu} b_i \frac{d\mathcal{N}}{dY} - (1 + \beta^{1-\nu} \mathcal{N}) \beta^\epsilon \frac{db_i}{dY} \\ = \frac{\beta^{-2\nu}}{\Omega} (1 + \beta^{1-\nu} \mathcal{N})^2 (\mathcal{N}_{upper}^2 - \mathcal{N}^2) \end{aligned} \quad (4.39)$$

$$\begin{aligned} \Rightarrow \beta^{\epsilon+3\nu-1} b_i \frac{d\mathcal{N}}{dY} - \beta^{\epsilon+3\nu-1} (\beta^{\nu-1} + \mathcal{N}) \frac{db_i}{dY} \\ = \frac{1}{\Omega} (\beta^{\nu-1} + \mathcal{N})^2 (\mathcal{N}_{upper}^2 - \mathcal{N}^2), \end{aligned} \quad (4.40)$$

which is the equation we solve within each boundary-layer.

#### 4.4.3 Scaling and Expanding the Square of the Upper Solution

Now we put the square of the upper solution into scaled and expanded form. Beginning with equation (4.32),

$$\frac{N_{upper}^2}{N_e^2} = 1 - \Omega \frac{db_i}{dy}, \quad (4.41)$$

and applying the scalings,

$$\beta^{-2\nu} \mathcal{N}_{upper}^2 = 1 - \Omega \beta^\epsilon \frac{db_i}{dY}. \quad (4.42)$$

The Taylor expansion of  $b_i$  about  $y = \xi$  is

$$b_i(Y) = b_i(\xi) + b_i'(\xi) \beta^{-\epsilon} Y + \frac{b_i''(\xi)}{2} \beta^{-2\epsilon} Y^2 + \frac{b_i'''(\xi)}{3!} \beta^{-3\epsilon} Y^3 + \dots$$

as stated in (4.27), and putting this into (4.42),

$$\beta^{-2\nu} \mathcal{N}_{upper}^2 = 1 - \Omega \beta^\epsilon \frac{d}{dY} \left( b_i(\xi) + b_i'(\xi) \beta^{-\epsilon} Y + \frac{b_i''(\xi)}{2} \beta^{-2\epsilon} Y^2 + \dots \right) \quad (4.43)$$

$$\Rightarrow \beta^{-2\nu} \mathcal{N}_{upper}^2 = 1 - \Omega b_i'(\xi) - \Omega b_i''(\xi) \beta^{-\epsilon} Y - \dots \quad (4.44)$$

It is a convenient shorthand to define

$$N_s = [N_{upper}]_{y=\xi}, \quad (4.45)$$

the scaled version of which is

$$\mathcal{N}_s = \beta^{2\nu} (1 - \Omega b'_i(\xi)) \quad (4.46)$$

by (4.41). Using this to further simplify equation (4.44),

$$\beta^{-2\nu} \mathcal{N}_{upper}^2 = \beta^{-2\nu} \mathcal{N}_s^2 - \Omega b''_i(\xi) \beta^{-\epsilon} Y - \dots \quad (4.47)$$

$$\Rightarrow \mathcal{N}_{upper}^2 = \mathcal{N}_s^2 - \Omega b''_i(\xi) \beta^{2\nu-\epsilon} Y - \frac{\Omega b'''_i(\xi)}{2} \beta^{2\nu-2\epsilon} Y^2 - \dots \quad (4.48)$$

Although the expansion of  $\mathcal{N}_{upper}^2$  is valid in both boundary-layers, the expansions of  $\mathcal{N}_{upper}$  and  $\mathcal{N}_{lower}$  change depending on which layer we are studying. Expansions of  $\mathcal{N}_{upper}$  and  $\mathcal{N}_{lower}$  are therefore deferred to later sections, where the two boundary-layers are treated individually.

#### 4.4.4 Determination of Scaling Parameters $\epsilon$ and $\nu$

The parameters  $\epsilon$  and  $\nu$  are determined by inspecting the scaled governing equation (4.40) and the expanded form of  $\mathcal{N}_{upper}^2$ . We expand  $\mathcal{N}$  as

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1 \beta^{-\epsilon} + \mathcal{N}_2 \beta^{-2\epsilon} + \dots,$$

noting that the greatest power of  $\beta$  in this expansion is zero. Assuming that  $\nu < 1$ , matching the greatest powers of  $\beta$  in equation (4.40) requires

$$\epsilon + 3\nu = 1, \quad (4.49)$$

provided that we also arrange for the greatest power of  $\beta$  in  $\mathcal{N}_{upper}^2$  to be zero.

When boundary-layer matching is performed between adjacent channels of upward and downward field-aligned current,  $\mathcal{N}_{upper}^2(y = \xi)$  is non-zero, so the greatest power of  $\beta$  on the right-hand side of (4.48) is automatically zero, removing any need to scale the density. We therefore satisfy (4.49) by taking,

$$\nu = 0, \quad \epsilon = 1. \quad (4.50)$$

When boundary-layer matching is performed at the edge of the depleted region farthest from the upward current-channel,  $\mathcal{N}_{upper}^2(y = \xi) = 0$ . This leaves  $\Omega b''_i(\xi) \beta^{2\nu-\epsilon} Y$  as the leading term on the right-hand side of (4.48); setting the power of  $\beta$  to zero,

$$2\nu - \epsilon = 0, \quad (4.51)$$

and solving this alongside (4.49) yields

$$\nu = \frac{1}{5}, \quad \epsilon = \frac{2}{5}. \quad (4.52)$$

Alternatively, one can obtain scalings for the two boundary-layers by inspection of  $\mathcal{N}_{upper}$  and  $\mathcal{N}_{lower}$ , requiring that the greatest power of  $\beta$  be zero in each expansion. The same values are obtained for  $\epsilon$  and  $\nu$ .

#### 4.4.5 Boundary-Layer Between Upward and Downward Current-Channels

##### 4.4.5.1 Preliminary

We begin piecing the global solution together, starting at the boundary-layer between the upward and downward current regions. This boundary-layer is sited at  $y = \xi$ , where the lower solution is singular and the upper solution is non-zero (because this side of the cavity has broadened). In Section 4.4.4 the scaling parameters for this layer were determined as  $\nu = 0$  and  $\epsilon = 1$ . It is useful to note that, since the lower solution is singular at  $y = \xi$ , equation (4.15) gives

$$\xi + c = \Omega b_i(\xi), \quad (4.53)$$

and meanwhile,

$$\mathcal{N}_s^2 = \frac{N_s^2}{N_e^2} = (1 - \Omega b_i'(\xi)) > 0; \quad (4.54)$$

both of these results are useful in the following algebra.

##### 4.4.5.2 Expansion of $\mathcal{N}_{upper}$

The expansion of  $\mathcal{N}_{upper}$  can be found by substituting

$$\mathcal{N}_{upper} = a_0 + a_1 \beta^{-1} Y + a_2 \beta^{-2} Y^2 + \dots \quad (4.55)$$

into

$$\mathcal{N}_{upper}^2 = \mathcal{N}_s^2 - \Omega b_i''(\xi) \beta^{-1} Y - \frac{\Omega b_i'''(\xi)}{2} \beta^{-2} Y^2 - \dots \quad (4.56)$$

which is obtained by substituting for  $\epsilon$  and  $\nu$  in equation (4.48). Doing so,

$$a_0^2 + 2a_0 a_1 \beta^{-1} Y + (2a_0 a_2 + a_1^2) \beta^{-2} Y^2 + \dots$$

$$= \mathcal{N}_s^2 - \Omega b_i''(\xi) \beta^{-1} Y - \frac{\Omega b_i'''(\xi)}{2} \beta^{-2} Y^2 - \dots \quad (4.57)$$

and we obtain the expansion coefficients,  $a_j$ , recursively by equating powers of  $Y$ :

$$a_0 = \mathcal{N}_s, \quad (4.58)$$

$$\begin{aligned} a_1 &= -\frac{\Omega b_i''(\xi)}{2a_0} \\ &= -\frac{\Omega b_i''(\xi)}{2\mathcal{N}_s}, \end{aligned} \quad (4.59)$$

$$\begin{aligned} a_2 &= \frac{1}{2a_0} (b_2 - a_1^2) \\ &= \frac{1}{2\mathcal{N}_s} \left( -\frac{\Omega b_i'''(\xi)}{2} - \left( \frac{\Omega b_i''(\xi)}{2\mathcal{N}_s} \right)^2 \right) \\ &= -\frac{\Omega}{4\mathcal{N}_s} \left( b_i'''(\xi) + \frac{\Omega b_i''^2(\xi)}{2\mathcal{N}_s^2} \right), \end{aligned} \quad (4.60)$$

and so on. Thus,

$$\mathcal{N}_{upper} = \mathcal{N}_s - \frac{\Omega b_i''(\xi)}{2\mathcal{N}_s \beta} Y - \frac{\Omega}{4\mathcal{N}_s \beta^2} \left( b_i'''(\xi) + \frac{\Omega b_i''^2(\xi)}{2\mathcal{N}_s^2} \right) Y^2 - \dots \quad (4.61)$$

#### 4.4.5.3 Expansion of $\mathcal{N}_{lower}$

Moving onto the lower solution, equation (4.13) states

$$\begin{aligned} \beta \frac{\mathcal{N}_{lower}}{N_e} &= \frac{y + c}{\Omega b_i - (y + c)} \\ \Rightarrow \beta \frac{\mathcal{N}_{lower}}{N_e} &= \frac{(y - \xi) + (\xi + c)}{\Omega b_i - (y - \xi) - (\xi + c)}, \end{aligned} \quad (4.62)$$

which becomes

$$\beta \mathcal{N}_{lower} = -\frac{\beta^{-1} Y + (\xi + c)}{-\Omega b_i + \beta^{-1} Y + (\xi + c)} \quad (4.63)$$

under the scalings. We have previously noted that  $\xi + c = \Omega b_i(\xi)$  (4.53), so

$$\beta \mathcal{N}_{lower} = -(\beta^{-1} Y + \Omega b_i(\xi)) (-\Omega b_i + \beta^{-1} Y + \Omega b_i(\xi))^{-1}. \quad (4.64)$$

Expanding  $b_i$  with the Taylor series given in (4.27) and performing a little algebra,

$$\beta \mathcal{N}_{lower} = -(\beta^{-1} Y + \Omega b_i(\xi)) \left( -\Omega \left( b_i(\xi) + b_i'(\xi) \beta^{-1} Y + \frac{b_i''(\xi)}{2} \beta^{-2} Y^2 + \dots \right) + (\beta^{-1} Y + \Omega b_i(\xi)) \right)^{-1}$$

$$\begin{aligned}
&= -(\beta^{-1}Y + \Omega b_i(\xi)) \left( (1 - \Omega b'_i(\xi))\beta^{-1}Y - \frac{\Omega b''_i(\xi)}{2}\beta^{-2}Y^2 + \dots \right)^{-1} \\
&= -(\beta^{-1}Y + \Omega b_i(\xi)) \left( \mathcal{N}_s^2\beta^{-1}Y - \frac{\Omega b''_i(\xi)}{2}\beta^{-2}Y^2 + \dots \right)^{-1}. \tag{4.65}
\end{aligned}$$

Since  $\mathcal{N}_s^2 \neq 0$ , the leading term in the ‘denominator bracket’ can be factored out, so

$$\begin{aligned}
\beta \mathcal{N}_{lower} &= -(\beta^{-1}Y + \Omega b_i(\xi)) \frac{\beta}{\mathcal{N}_s^2 Y} \left( 1 - \frac{\Omega b''_i(\xi)}{2\mathcal{N}_s^2}\beta^{-1}Y + \dots \right)^{-1} \\
&= -(\beta^{-1}Y + \Omega b_i(\xi)) \frac{\beta}{\mathcal{N}_s^2 Y} \left( 1 + \frac{\Omega b''_i(\xi)}{2\mathcal{N}_s^2}\beta^{-1}Y + \dots \right) \\
\Rightarrow \mathcal{N}_{lower} &= -\left( \frac{\Omega b_i(\xi)}{\mathcal{N}_s^2 Y} + \frac{1}{\beta}\mathcal{N}_s^2 \right) \left( 1 + \frac{\Omega b''_i(\xi)}{2\mathcal{N}_s^2}\beta^{-1}Y + \dots \right),
\end{aligned}$$

and multiplying out the brackets and collecting terms in  $Y$ , the final result is

$$\mathcal{N}_{lower} = -\frac{\Omega b_i(\xi)}{\mathcal{N}_s^2 Y} - \frac{1}{\beta}\mathcal{N}_s^2 - \frac{\Omega^2 b_i(\xi)b''_i(\xi)}{2\mathcal{N}_s^2} + \dots \tag{4.66}$$

#### 4.4.5.4 Behaviour of $\mathcal{N}_{upper}$ and $\mathcal{N}_{lower}$ as $Y \rightarrow 0$

By inspection of (4.61) and (4.66), as  $Y \rightarrow 0$  the upper solution tends to the constant  $\mathcal{N}_s$ , whilst the lower solution is singular in this limit and behaves as  $-\Omega b_i(\xi)/\mathcal{N}_s^2 Y$ . These outer solutions contain at most one singular term, so solving for  $\mathcal{N}_0$  (as introduced in (4.28)) will be sufficient to build a global solution.

#### 4.4.5.5 Solution Inside Boundary-layer

Putting  $\nu = 0$  and  $\epsilon = 1$  into the scaled governing equation (4.40), we wish to find an approximate solution to

$$b_i \frac{d\mathcal{N}}{dY} - (\beta^{-1} + \mathcal{N}) \frac{db_i}{dY} = \frac{1}{\Omega} (\beta^{-1} + \mathcal{N})^2 (\mathcal{N}_{upper}^2 - \mathcal{N}^2). \tag{4.67}$$

Expanding  $b_i$  and  $\mathcal{N}$  as given in equations (4.27) and (4.28), substituting for  $\mathcal{N}_{upper}^2$  with (4.56), and equating terms in  $\beta^0$  gives

$$b_i(\xi) \frac{d\mathcal{N}_0}{dY} = \frac{1}{\Omega} \mathcal{N}_0^2 (\mathcal{N}_s^2 - \mathcal{N}_0^2), \tag{4.68}$$

which is a separable ordinary differential equation for  $\mathcal{N}_0$ . Rearranging and using partial fractions leads to the solution

$$\frac{Y}{\Omega b_i(\xi)} = -\frac{1}{\mathcal{N}_s^2 \mathcal{N}_0} - \frac{1}{2\mathcal{N}_s^3} \ln(\mathcal{N}_s - \mathcal{N}_0) + \frac{1}{2\mathcal{N}_s^3} \ln(\mathcal{N}_s + \mathcal{N}_0). \quad (4.69)$$

Strictly, the general solution to (4.68) contains a constant of integration, however experience shows that it has at most a small effect on  $\mathcal{N}_0$ , so we neglect it with little error.

#### 4.4.5.6 Asymptotic Behaviour of $\mathcal{N}_0$

Having obtained an implicit expression for  $\mathcal{N}_0$ , we check its asymptotic behaviour. As  $|Y| \rightarrow \infty$  in the direction of the upper steady-state, the left-hand side of (4.69) goes to infinity (in this limit  $Y$  and  $b_i(\xi)$  have the same sign). This must be balanced by the second term on the right-hand side, giving

$$\begin{aligned} \frac{Y}{\Omega b_i(\xi)} &\approx -\frac{1}{2\mathcal{N}_s^3} \ln(\mathcal{N}_s - \mathcal{N}_0) \\ \Rightarrow \mathcal{N}_0 &\approx \mathcal{N}_s - \exp\left(-\frac{2\mathcal{N}_s^3 Y}{\Omega b_i(\xi)}\right). \end{aligned} \quad (4.70)$$

Hence,  $\mathcal{N}_0 \rightarrow \mathcal{N}_s$  as  $Y \rightarrow \text{sign}(b_i(\xi))\infty$ , which matches the behaviour of the upper solution (4.61) as  $Y \rightarrow 0$ . This allows us to construct a solution on the side of  $\xi$  where the outer solution is the upper steady-state.

As  $|Y| \rightarrow \infty$  in the direction of the lower steady-state, the left-hand side of (4.69) goes to minus infinity. This can only be balanced by the first term on the right-hand side, so, in this limit,

$$\begin{aligned} \frac{Y}{\Omega b_i(\xi)} &\approx -\frac{1}{\mathcal{N}_s^2 \mathcal{N}_0} \\ \Rightarrow \mathcal{N}_0 &\approx -\frac{\Omega b_i(\xi)}{\mathcal{N}_s^2 Y}. \end{aligned} \quad (4.71)$$

Thus, the behaviour of  $\mathcal{N}_0$  in this limit matches the behaviour of the lower solution as  $Y \rightarrow 0$ . This allows us to construct a solution on the side of  $\xi$  where the outer solution is the lower steady-state.

#### 4.4.5.7 Steady-state Constructed Across the Boundary-layer

Now we are ready to construct the steady-state across the boundary-layer. Taking  $\mathcal{N}_0$  as the inner solution, we set

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_{upper} - \mathcal{N}_s, \quad (4.72)$$

on the upper steady-state side of  $\xi$ , and

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_{lower} + \frac{\Omega b_i(\xi)}{\mathcal{N}_s^2 Y}, \quad (4.73)$$

on the lower steady-state side of  $\xi$ . Here  $\mathcal{N}_0$  is given implicitly by (4.69) above.

#### 4.4.5.8 Result in Unscaled Form

To finish, the unscaled solution is

$$N = N_0 + N_{upper} - N_s, \quad (4.74)$$

on the upper steady-state side of  $\xi$  and

$$N = N_0 + N_{lower} + \frac{N_e^3 \Omega b_i(\xi)}{\beta N_s^2 (y - \xi)}, \quad (4.75)$$

on the lower steady-state side of  $\xi$ , where

$$N_s = N_{upper}(\xi) \quad (4.76)$$

and  $N_0$  is given implicitly by,

$$\begin{aligned} -\frac{\beta(y - \xi)}{N_e^3 \Omega b_i(\xi)} &= \frac{1}{N_s^2 N_0} + \frac{1}{2N_s^3} \ln(N_s - N_0) \\ &\quad - \frac{1}{2N_s^3} \ln(N_s + N_0). \end{aligned} \quad (4.77)$$

### 4.4.6 Boundary-Layer Farthest From Upward Current-Channel

#### 4.4.6.1 Preliminary

Boundary-layer matching follows the same process farthest from the upward current-channel and as between the upward and downward current-channel; nonetheless, there are some important differences in the detail because, this time, the upper solution is zero at  $y = \xi$ . In Section 4.4.4 the scaling parameters were determined to be  $\nu = 1/5$  and  $\epsilon = 2/5$ . The lower solution is singular at  $y = \xi$ , so we still have

$$\xi + c = \Omega b_i(\xi), \quad (4.78)$$

but this time,

$$1 - \Omega b'_i(\xi) = \frac{N_s^2}{N_e^2} = \beta^{-2/5} \mathcal{N}_s^2 = 0. \quad (4.79)$$

#### 4.4.6.2 Expansion of $\mathcal{N}_{upper}$

Substituting for  $\epsilon$  and  $\nu$  in equation (4.48),

$$\mathcal{N}_{upper}^2 = -\Omega b''_i(\xi)Y - \frac{\Omega b'''_i(\xi)}{2}\beta^{-2/5}Y^2 - \dots \quad (4.80)$$

This time, we can obtain the expansion of  $\mathcal{N}_{upper}$  by substituting

$$\mathcal{N}_{upper} = Y^{1/2} \left( a_0 + a_1 \beta^{-2/5} Y + \dots \right) \quad (4.81)$$

into the left-hand side of (4.80) to get

$$a_0^2 Y + 2a_0 a_1 \beta^{-2/5} Y^2 + \dots = -\Omega b''_i(\xi)Y - \frac{\Omega b'''_i(\xi)}{2}\beta^{-2/5}Y^2 - \dots \quad (4.82)$$

Equating powers of  $Y$ , the expansion coefficients,  $a_j$ , are obtained recursively as:

$$a_0 = \sqrt{-\Omega b''_i(\xi)}, \quad (4.83)$$

$$\begin{aligned} a_1 &= \frac{1}{2a_0} \left( -\frac{\Omega b'''_i(\xi)}{2} \right) \\ &= -\frac{\Omega b'''_i(\xi)}{4\sqrt{-\Omega b''_i(\xi)}} \\ &= -\frac{b'''_i(\xi)\sqrt{-\Omega b''_i(\xi)}}{4b''_i(\xi)}, \end{aligned} \quad (4.84)$$

and so on. Thus,

$$\mathcal{N}_{upper} = \sqrt{-\Omega b''_i(\xi)}Y - \frac{\beta^{-2/5} b'''_i(\xi)\sqrt{-\Omega b''_i(\xi)}YY}{4b''_i(\xi)} - \dots \quad (4.85)$$

#### 4.4.6.3 Expansion of $\mathcal{N}_{lower}$

The lower solution satisfies equation (4.62) as before, so that

$$\beta \frac{N_{lower}}{N_e} = \frac{(y - \xi) + (\xi + c)}{\Omega b_i - (y - \xi) - (\xi + c)},$$

which becomes

$$\beta^{4/5} \mathcal{N}_{lower} = -\frac{\beta^{-2/5} Y + (\xi + c)}{-\Omega b_i + \beta^{-2/5} Y + (\xi + c)} \quad (4.86)$$

under the new scalings. Once again,  $\xi + c = \Omega b_i(\xi)$  (4.78) so

$$\beta^{4/5} \mathcal{N}_{lower} = -\frac{\beta^{-2/5} Y + \Omega b_i(\xi)}{-\Omega b_i + \beta^{-2/5} Y + \Omega b_i(\xi)}. \quad (4.87)$$

Expanding  $b_i$  with the Taylor series given in (4.27) and performing a little algebra,

$$\begin{aligned} \beta^{4/5} \mathcal{N}_{lower} &= -\left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left( \frac{-\Omega (b_i(\xi) + b'_i(\xi) \beta^{-2/5} Y + \dots)}{+\beta^{-2/5} Y + \Omega b_i(\xi)} \right)^{-1} \\ &= -\left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left( \frac{(1 - \Omega b'_i(\xi)) \beta^{-2/5} Y - \frac{\Omega b''_i(\xi)}{2} \beta^{-4/5} Y^2}{-\frac{\Omega b'''_i(\xi)}{3!} \beta^{-6/5} Y^3 + \dots} \right)^{-1} \\ &= -\left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left( \frac{\mathcal{N}_s^2 \beta^{-2/5} Y - \frac{\Omega b''_i(\xi)}{2} \beta^{-4/5} Y^2}{-\frac{\Omega b'''_i(\xi)}{3!} \beta^{-6/5} Y^3 + \dots} \right)^{-1} \\ &= \frac{\beta^{4/5}}{\Omega} \left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left( \frac{b''_i(\xi)}{2} Y^2 + \frac{b'''_i(\xi)}{3!} \beta^{-2/5} Y^3 + \dots \right)^{-1}. \quad (4.88) \end{aligned}$$

If  $b''_i(\xi) \neq 0$ , then the leading term in the ‘denominator bracket’ can be factored out, so

$$\begin{aligned} \mathcal{N}_{lower} &= \frac{2}{\Omega b''_i(\xi) Y^2} \left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left(1 + \frac{b'''_i(\xi)}{3b''_i(\xi)} \beta^{-2/5} Y + \dots\right)^{-1} \\ &= \frac{2}{\Omega b''_i(\xi) Y^2} \left(\beta^{-2/5} Y + \Omega b_i(\xi)\right) \left(1 - \frac{b'''_i(\xi)}{3b''_i(\xi)} \beta^{-2/5} Y + \dots\right) \\ &= \left(\frac{2b_i(\xi)}{b''_i(\xi) Y^2} + \frac{2b'_i \beta^{-2/5}}{\Omega b''_i(\xi) Y}\right) \left(1 - \frac{b'''_i(\xi)}{3b''_i(\xi)} \beta^{-2/5} Y + \dots\right) \end{aligned}$$

Equation (4.79) gives  $\Omega = 1/b'_i(\xi)$ , which removes  $\Omega$ :

$$\mathcal{N}_{lower} = \left(\frac{2b_i(\xi)}{b''_i(\xi) Y^2} + \frac{2b'_i b'_i(\xi) \beta^{-2/5}}{b''_i(\xi) Y}\right) \left(1 - \frac{b'''_i(\xi)}{3b''_i(\xi)} \beta^{-2/5} Y + \dots\right)$$

Finally, multiplying the brackets and collecting terms in  $Y$ , the scaled and expanded lower solution is

$$\mathcal{N}_{lower} = \frac{2b_i(\xi)}{b''_i(\xi) Y^2} + \left(\frac{2b'_i(\xi)}{b''_i(\xi)} - \frac{2b_i(\xi) b'''_i(\xi)}{3b''_i{}^2(\xi)}\right) \frac{\beta^{-2/5}}{Y} + \dots \quad (4.89)$$

#### 4.4.6.4 Behaviour of $\mathcal{N}_{upper}$ and $\mathcal{N}_{lower}$ as $Y \rightarrow 0$

By inspection, as  $Y \rightarrow 0$ , the upper solution tends to zero, and the lower solution is singular with terms behaving as  $1/Y$  and  $1/Y^2$ . Because outer solutions involve two singular terms, we solve for  $\mathcal{N}_0$  and  $\mathcal{N}_1$  to construct a global solution.

#### 4.4.6.5 Solution Inside Boundary-layer

Putting  $\nu = 1/5$  and  $\epsilon = 2/5$  into the scaled governing equation (4.40), we wish to find an approximate solution to

$$b_i \frac{d\mathcal{N}}{dY} - \left( \beta^{-4/5} + \mathcal{N} \right) \frac{db_i}{dY} = \frac{1}{\Omega} \left( \beta^{-4/5} + \mathcal{N} \right)^2 \left( \mathcal{N}_{upper}^2 - \mathcal{N}^2 \right). \quad (4.90)$$

Expanding  $b_i$  and  $\mathcal{N}$  as given in equations (4.27) and (4.28), substituting for  $\mathcal{N}_{upper}^2$  with (4.80), and equating terms in  $\beta^0$  gives

$$\frac{d\mathcal{N}_0}{dY} = -\frac{\mathcal{N}_0^4}{\Omega b_i(\xi)} - \frac{b_i''(\xi)}{b_i(\xi)} \mathcal{N}_0^2 Y, \quad (4.91)$$

and equating terms in  $\beta^{-2/5}$  gives

$$\frac{d\mathcal{N}_1}{dY} = \mathcal{N}_0 \frac{b_i'(\xi)}{b_i(\xi)} (1 - 4\mathcal{N}_0^2 \mathcal{N}_1) - \left( 2\mathcal{N}_0 \mathcal{N}_1 \frac{b_i''(\xi)}{b_i(\xi)} + \frac{d\mathcal{N}_0}{dY} \frac{b_i'(\xi)}{b_i(\xi)} \right) Y - \frac{\mathcal{N}_0^2}{2} \frac{b_i'''(\xi)}{b_i(\xi)} Y^2. \quad (4.92)$$

Equation (4.92) is simplified by using equation (4.91) to remove  $d\mathcal{N}_0/dY$ , giving

$$\begin{aligned} \frac{d\mathcal{N}_1}{dY} &= \mathcal{N}_0 \frac{b_i'(\xi)}{b_i(\xi)} (1 - 4\mathcal{N}_0^2 \mathcal{N}_1) \\ &\quad + \frac{\mathcal{N}_0}{b_i(\xi)} \left( \frac{\mathcal{N}_0^3 b_i'^2(\xi)}{b_i(\xi)} - 2\mathcal{N}_1 b_i''(\xi) \right) Y \\ &\quad + \frac{\mathcal{N}_0^2}{b_i(\xi)} \left( \frac{b_i'(\xi) b_i''(\xi)}{b_i(\xi)} - \frac{b_i'''(\xi)}{2} \right) Y^2. \end{aligned} \quad (4.93)$$

I am, as yet, unaware of a general solution to ordinary differential equations (4.91) and (4.93), however, it is quite possible to proceed with  $\mathcal{N}_0$  and  $\mathcal{N}_1$  given implicitly by these equations.

#### 4.4.6.6 Asymptotic Behaviour of $\mathcal{N}_0$ and $\mathcal{N}_1$

It is valid to truncate the inner solution at  $\mathcal{N}_{inner} = \mathcal{N}_0 + \beta^{-2/5} \mathcal{N}_1$ , provided this solution picks up any singularities in the outer solution. We must therefore establish the asymptotic behaviour

of  $\mathcal{N}_0$  and  $\mathcal{N}_1$ .

First, let us examine the asymptotic behaviour of  $\mathcal{N}_0$ . As  $|Y| \rightarrow \infty$  in the direction of the upper steady-state, the second term on the right-hand side of (4.91) goes to  $\text{sign}(b_i)\infty$ . This is balanced by the first term on the right-hand side, so, in this limit,

$$\begin{aligned} \frac{\mathcal{N}_0^4}{\Omega b_i(\xi)} &\approx -\frac{b_i''(\xi)}{b_i(\xi)} \mathcal{N}_0^2 Y \\ \Rightarrow \mathcal{N}_0 &\approx \sqrt{-\Omega b_i''(\xi) Y}. \end{aligned} \quad (4.94)$$

As  $|Y| \rightarrow \infty$  in the direction of the lower steady-state, the second term on the right-hand side of (4.91) goes to  $-\text{sign}(b_i)\infty$ . This is balanced by the term on the left-hand side, giving

$$\begin{aligned} \frac{d\mathcal{N}_0}{dY} &\approx -\frac{b_i''(\xi)}{b_i(\xi)} \mathcal{N}_0^2 Y \\ \Rightarrow -\int \frac{d\mathcal{N}_0}{\mathcal{N}_0^2} &\approx \frac{b_i''(\xi)}{b_i(\xi)} \int Y dY \\ \Rightarrow \mathcal{N}_0 &\approx \frac{2b_i(\xi)}{b_i''(\xi) Y}. \end{aligned} \quad (4.95)$$

Next we consider the asymptotic behaviour of  $\mathcal{N}_1$ . As  $|Y| \rightarrow \infty$  in the direction of the upper steady-state, we can use (4.94) to substitute for  $\mathcal{N}_0$  in (4.93). After rearranging, this gives

$$\frac{d\mathcal{N}_1}{dY} \approx \frac{b_i'(\xi)}{b_i(\xi)} \sqrt{-\Omega b_i''(\xi) Y} + \frac{2\mathcal{N}_1 b_i''(\xi)}{b_i(\xi)} \sqrt{-\Omega b_i''(\xi) Y} + \frac{\Omega b_i''(\xi) b_i'''(\xi)}{2b_i(\xi)} Y^3. \quad (4.96)$$

In this limit, the term in  $Y^3$  is balanced by the term in  $\sqrt{-\Omega b_i''(\xi) Y}$ , so

$$\begin{aligned} \frac{2\mathcal{N}_1 b_i''(\xi)}{b_i(\xi)} \sqrt{-\Omega b_i''(\xi) Y} &\approx -\frac{\Omega b_i''(\xi) b_i'''(\xi)}{2b_i(\xi)} Y^3 \\ \Rightarrow \mathcal{N}_1 &\approx -\frac{b_i'''(\xi) \sqrt{-\Omega b_i''(\xi) Y}}{4b_i''(\xi)}. \end{aligned} \quad (4.97)$$

As  $|Y| \rightarrow \infty$  in the direction of the lower steady-state, we can use (4.95) to substitute for  $\mathcal{N}_0$  in (4.93). After rearranging, this gives

$$\frac{d\mathcal{N}_1}{dY} \approx -\frac{4\mathcal{N}_1}{Y} + \frac{6b_i'(\xi)}{b_i''(\xi) Y^2} - \frac{2b_i(\xi) b_i'''(\xi)}{b_i''^2(\xi) Y^2} - \frac{32\mathcal{N}_1 b_i^2(\xi) b_i'(\xi)}{b_i''(\xi) Y^6} + \frac{16b_i^2(\xi) b_i'(\xi)}{b_i''^4(\xi) Y^7}. \quad (4.98)$$

In this limit, terms in  $Y^{-6}$  and  $Y^{-7}$  may be neglected, so we solve the following first order ordinary differential equation:

$$\frac{d\mathcal{N}_1}{dY} + \frac{4}{Y} \mathcal{N}_1 \approx \left( \frac{6b_i'(\xi)}{b_i''(\xi)} - \frac{2b_i(\xi) b_i'''(\xi)}{b_i''^2(\xi)} \right) \frac{1}{Y^2}$$

$$\Rightarrow \mathcal{N}_1 \approx \left( \frac{2b'_i(\xi)}{b''_i(\xi)} - \frac{2b_i(\xi)b'''_i(\xi)}{3b''_i{}^2(\xi)} \right) \frac{1}{Y}. \quad (4.99)$$

From these asymptotic solutions we see that

$$\mathcal{N}_{inner} = \mathcal{N}_0 + \beta^{-2/5} \mathcal{N}_1$$

picks up the behaviour of the outer solutions in the appropriate limits, making this a valid truncation for the inner solution.

#### 4.4.6.7 Steady-state Constructed Across Boundary-layer

Since

$$\mathcal{N}_{inner} = \mathcal{N}_0 + \beta^{-2/5} \mathcal{N}_1$$

has the desired asymptotic behaviour, we construct the complete steady-state as

$$\mathcal{N} = \mathcal{N}_0 + \beta^{-2/5} \mathcal{N}_1 + \mathcal{N}_{upper} - \sqrt{-\Omega b''_i(\xi)Y} + \frac{\beta^{-2/5} b'''_i(\xi) \sqrt{-\Omega b''_i(\xi)Y} Y}{4b''_i(\xi)}, \quad (4.100)$$

on the upper steady-state side of  $\xi$ , and

$$\mathcal{N} = \mathcal{N}_0 + \beta^{-2/5} \mathcal{N}_1 + \mathcal{N}_{lower} - \frac{2b_i(\xi)}{b''_i(\xi)Y^2} + \left( \frac{2b'_i(\xi)}{b''_i(\xi)} - \frac{2b_i(\xi)b'''_i(\xi)}{3b''_i{}^2(\xi)} \right) \frac{\beta^{-2/5}}{Y}, \quad (4.101)$$

on the lower steady-state side of  $\xi$ . Here  $\mathcal{N}_0$  and  $\mathcal{N}_1$  are the solutions to ordinary differential equations (4.91) and (4.93) respectively.

#### 4.4.6.8 Result in Unscaled Form

To finish, the unscaled solution is

$$N = N_0 + \beta^{-2/5} N_1 + N_{upper} - N_e \sqrt{-\Omega b''_i(\xi)(y - \xi)} + \frac{N_e b'''_i(\xi) \sqrt{-\Omega b''_i(\xi)(y - \xi)}(y - \xi)}{4b''_i(\xi)}, \quad (4.102)$$

on the upper steady-state side of  $\xi$  and

$$N = N_0 + \beta^{-2/5} N_1 + N_{lower} - \frac{2N_e b_i(\xi) \beta^{-1}}{b''_i(\xi)(y - \xi)^2} + \left( \frac{2b'_i(\xi)}{b''_i(\xi)} - \frac{2b_i(\xi)b'''_i(\xi)}{3b''_i{}^2(\xi)} \right) \frac{N_e \beta^{-1}}{(y - \xi)}, \quad (4.103)$$

on the lower steady-state side of  $\xi$ , where  $N_0$  and  $N_1$  are the solutions of the following ordinary differential equations:

$$\frac{dN_0}{dy} = -\frac{\beta N_0}{\Omega b_i(\xi)} + \frac{\beta b_i''(\xi)}{N_e b_i(\xi)} N_0^2 (y - \xi), \quad (4.104)$$

$$\begin{aligned} \frac{dN_1}{dy} = & \frac{N_0}{\Omega b_i(\xi)} \left( \beta^{2/5} - \frac{\beta N_0^2 N_1}{N_e^3} \right) \\ & + \frac{N_0}{N_e b_i(\xi)} \left( \frac{\beta^{7/5} N_0^3}{N_e^2 \Omega^2 b_i(\xi)} - 2\beta N_1 b_i''(\xi) \right) (y - \xi) \\ & + \frac{\beta^{7/5} N_0^2}{N_e b_i(\xi)} \left( \frac{b_i''(\xi)}{\Omega b_i(\xi)} - \frac{b_i'''(\xi)}{2} \right) (y - \xi)^2. \end{aligned} \quad (4.105)$$

These can be solved numerically if we note that the asymptotic behaviour

$$N_0 \approx -N_e \sqrt{\Omega b_i''(\xi)(y - \xi)}, \quad (4.106)$$

$$N_1 \approx -\frac{N_e \beta^{2/5} b_i'''(\xi) \sqrt{\Omega b_i''(\xi)(y - \xi)}(y - \xi)}{4b_i''(\xi)}, \quad (4.107)$$

as  $|y| \rightarrow \infty$  in the direction of the upper steady-state, provides boundary-conditions for numerical integration.

#### 4.4.7 Global Solution

The boundary-layer matching has now been completed in two distinct stages, but we would like to combine these separate solutions into a single global solution. Fortunately, this is straightforward. If boundary-layer matching between the upward and downward channels is performed about  $y = a$ , and boundary-layer matching far from the upward channel is performed about  $y = b$ , then the complete solution can be constructed using (4.74) and (4.102) outside the plasma-density cavity, and taking

$$\begin{aligned} N = & N_{lower} + N_{a0} + N_{b0} + \beta^{-2/5} N_{b1} \\ & + \frac{N_e^3 \Omega b_i(a)}{\beta N_s^2 (y - a)} - \frac{2N_e b_i(b) \beta^{-1}}{b_i''(b)(y - b)^2} \\ & + \left( \frac{2b_i'(b)}{b_i''(b)} - \frac{2b_i(b)b_i'''(b)}{3b_i''(b)} \right) \frac{N_e \beta^{-1}}{(y - b)} \end{aligned} \quad (4.108)$$

inside the plasma-density cavity. Here,  $N_{a0}$  is the solution to (4.77) with  $\xi = a$ ,  $N_{b0}$  is the solution to (4.104) with  $\xi = b$ , and  $N_{b1}$  is the solution to (4.105) with  $\xi = b$ .

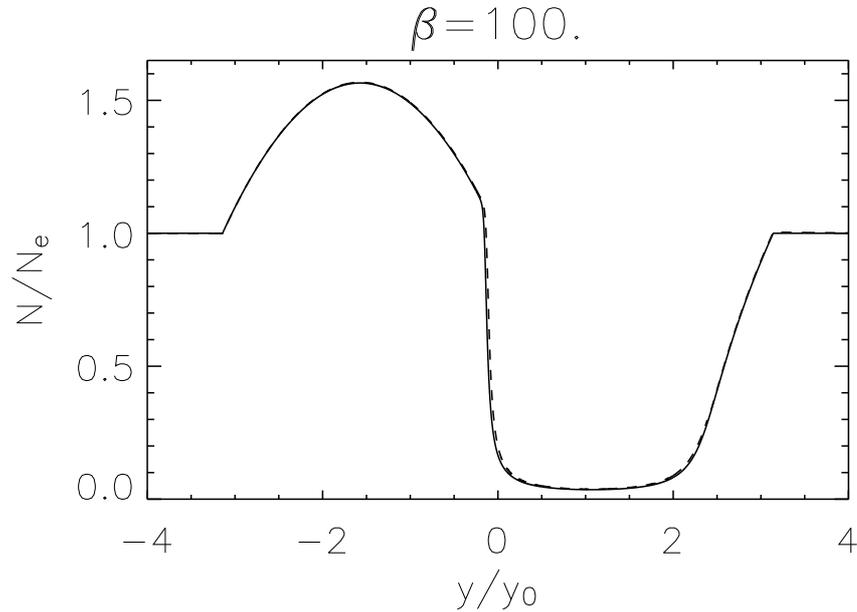


Figure 4.7: Comparison of steady-states obtained numerically (solid line) and analytically (dashed line) for  $\beta = 100$ . Reproduced from Russell et al. (2010) by permission of American Geophysical Union. Copyright 2010 American Geophysical Union.

#### 4.4.8 Comparison with Numerical Steady-States

The analytic solution agrees very well with numerical steady-states: a comparison is shown for  $\beta = 100$  in Figure 4.7 and the agreement is excellent. For all simulations with  $\beta \geq 20$ , the area between the two curves is less than 3.7% of the area under the numerical solution, with the best agreement obtained for large  $\beta$ . For very low values of  $\beta$ , agreement can be improved by including higher order corrections in the analytic solution.

#### 4.4.9 Shortest Length-Scale in Steady-State

As a last comment on the analytic solution, observe that the scaling of  $Y$  during boundary-layer matching captures the width of each boundary-layer: if the  $y$ -coordinate scales as  $Y = \beta^\epsilon(y - \xi)$  for the analysis, then the width of the boundary-layer is proportional to  $\beta^{-\epsilon}$ . Referring to the scaling obtained in Section 4.4.2, the width of the boundary-layer next to the upward current-channel is proportional to  $\beta^{-1}$ , and the width of the boundary-layer farthest from the upward current-channel is proportional to  $\beta^{-2/5}$ . It follows that the shortest length-scale in our steady-state is proportional to  $1/\beta$ , and it occurs between the upward and downward current-channels. This agrees with the numerical results presented in Section 4.2.4.

## 4.5 Discussion

We have obtained self-consistent steady-states for a sheet E-region coupled to an ideal magnetosphere by field-aligned currents, using both computer simulations and analytic methods. These methods are applicable to any current system driven by an incident Alfvén wave, and illustrate the features of such a system at late times. Analytic results are particularly useful: they agree well with computer simulations, and have provided insight into the formation and broadening of density cavities, their width, the minimum density, and the shortest length-scale in the steady state.

These steady-states can be viewed in terms of the need to balance gains and losses of electrons in the E-region. Three processes are involved (see terms in equation (3.2)): ionisation produces electrons at a rate which remains constant on the time-scale of our model (minutes or tens of minutes); recombination is a loss process, the rate of which is proportional to  $N^2$ ; and field-aligned currents act as a gain term for upward current and a loss term for downward current. Note that for a highly reflective E-region, where the reflection coefficient does not vary significantly from  $r = -1$ , the total current density drawn by the magnetospheric current is  $j_z = 2j_i$ .

In the absence of field-aligned current, ionisation and recombination balance one another, producing a uniform E-region. When upward field-aligned current is added, gains initially exceed losses, so number density is increased by upward field-aligned current. This, in turn, increases the recombination rate until the loss of E-region electrons matches electron gains (which remain fixed throughout this process). Thus, because recombination naturally increases to balance upward field-aligned currents, all upward currents are easily sustained by the E-region.

If we add downward current to the top of a uniform E-region, then we increase the loss of E-region electrons. Here, ionisation is the only source of E-region electrons, and this produces electrons at a fixed rate. Hence, if high reflectivity with  $r \approx -1$  is to be maintained, then recombination must be reduced until the total losses to downward current and recombination matches gains from ionisation. If, however, the absolute downward current density drawn from a highly reflective E-region exceeds  $j_c = \alpha n_e^2 h e$ , then ionisation cannot sustainably support the downward current, even if recombination goes to zero. In such a case, the reflection coefficient must change significantly from  $r = -1$  to reduce the total downward current density. These modifications correspond to the formation and widening of an E-region plasma-density cavity, with associated broadening of the downward current channel, and must take place if the absolute total current density (incident plus reflected, which is  $|2j_i|$  for  $r = -1$ ) exceeds  $j_c$  in the downward current region. Thus, if we know the extreme downward current density in the incident Alfvén wave, we can determine whether or not an E-region density cavity will form and widen, with associated broadening of the downward current.

It follows that downward current systems can be classified into two types, according to the

ratio of the extreme downward current-density in the incident Alfvén wave to a critical current density,  $j_c$ , determined by ionospheric parameters. In the first regime,  $|2j_i| < j_c$  throughout all downward current regions, so the current density is weak enough that the ionospheric reflection coefficient is not significantly altered from  $r = -1$ : in this case, either half of the ionosphere-magnetosphere system could be solved by treating the other part passively, with  $r = -1$  as the boundary-condition. If, however, the current density is sufficiently strong that  $|2j_i| > j_c$  somewhere in the downward current channel, then the height-integrated E-region number density is significantly suppressed from its initial value: a density cavity forms (with  $r > -1$ ) and the downward current channel widens as the cavity broadens. This classification was previously noted by Cran-McGreehin et al. (2007) for downward current regions ( $|2j_i/j_c| \equiv \mathcal{W}$  in their study), but their result was limited by the assumption of a tanh profile for the incident magnetic field perturbation, and the assumption (in their analytic work) that E-region plasma-density is a step function that jumps from  $N = 0$  inside the density cavity to  $N = N_e$  outside the density cavity. The present work, which is free from such restrictions, is a powerful generalisation of this important result.

The concept of upper and lower steady-states is an important analytic tool, and we have had remarkable success in using these solutions to construct an accurate description of the global steady-state through boundary-layer matching. It is also interesting to note the existence of a degenerate region where both upper and lower steady-states are valid (neither solution contradicts its underlying assumptions): referring to Figure 4.5, the lower steady-state is valid between the intersections of  $f(y) = c$  and  $f(y) = \Omega b_i(y) - y$ , and the upper steady-state is valid everywhere except between the two turning points of  $f(y) = \Omega b_i(y) - y$ ; this leaves an interval in  $y$  where both solutions are valid, between the minimum turning point and the outlying intersection. Computer simulations have made it clear that the final steady-state in this degenerate region lies along the lower solution, but why should the lower solution be favoured? This curiosity is addressed in Chapter 6, where we will also see the importance of degenerate steady-states for the dynamics of this system.

The final feature that deserves further discussion, beyond that included in earlier sections of this chapter, is the shortest length-scale in the steady-state. We have demonstrated that the finest scale present in the steady state scales as  $1/\beta$ . Therefore, for given  $\eta$ ,  $\alpha$ ,  $h$  and  $N_e$  there is always a threshold value of  $\beta$  above which electron inertial effects should be considered in the magnetosphere when computing the steady-state. As an example, the runs presented in Sect. 4.2 exhibit a fine scale,  $l_{min} \approx 9.26y_0/\beta$  (4.6). In a study of field-line resonances, Wei et al. (1994) showed that electron inertial effects are important in the magnetosphere for  $\lambda \lesssim 6\lambda_e$ , where  $\lambda_e = \sqrt{m^-/\mu_0 n_m e^2}$  is the electron inertial length in the magnetosphere. Assuming the width of the current channel is approximately  $w = \pi y_0$ , electron inertial effects in the magnetosphere will

modify the steady-state for

$$\begin{aligned}
 6\lambda_e &\gtrsim l_{min} \approx \frac{9.26y_0}{\beta} \approx \frac{9.26w}{\pi\beta} \\
 \Rightarrow \beta &\gtrsim \frac{w}{2\lambda_e}.
 \end{aligned}
 \tag{4.109}$$

In order to establish the effect of electron inertial effects on the final steady-state, computer simulations were performed using the inertial code described in Section 3.5, under conditions similar to those used for the ideal simulations presented in Section 4.2. These showed that the ideal solution is an excellent description of the steady-state if the electron inertial length at the bottom of the magnetosphere,  $\lambda_e$ , is less than the shortest length-scale in the ideal steady-state,  $l_{min}$ . If, however,  $l_{min} \lesssim \lambda_e$ , then the steep gradient between the upward and downward current channels is smoothed out, so that the shortest length-scale is approximately  $\lambda_e$ .

# Existence and Nature of Ionosphere-Magnetosphere Waves

---

## 5.1 Introduction

Perhaps the principal result of this thesis is the discovery of a new type of wave which arises from coupling between the ionosphere and magnetosphere. For the time being, I have simply called these ionosphere-magnetosphere waves, or IM-waves for short. There is much that could be done in future to expand the current theory - I believe it to be virgin territory - but the fundamentals have now been established and are presented here for the first time.

Although the idea of ionosphere-magnetosphere waves is new, there is a closely connected body of existing work on ‘ionospheric feedback instability’ (IFI). I will, therefore, also highlight some of the connections between IM-waves and IFI, hoping that a useful transfer of knowledge can accelerate our understanding of IM-waves in years to come.

## 5.2 Ideal IM-Waves

### 5.2.1 Advection Equation

Interactions between an ideal magnetosphere and a thin ‘sheet’ E-region, are characterised by a single governing equation (3.46) that states

$$\frac{\partial N}{\partial t} + \frac{2}{\mu_0 e} \frac{\partial}{\partial y} \left( \frac{\beta N b_i}{N_e + \beta N} \right) = \frac{\alpha}{h} (N_e^2 - N^2).$$

Expanding the  $y$ -derivative, this can be written in the form of an *advection* equation:

$$\frac{\partial N}{\partial t} + v_{IM} \frac{\partial N}{\partial y} = F, \tag{5.1}$$

where the characteristic velocity

$$v_{IM} = \frac{2\beta N_e b_i}{e\mu_0 (N_e + \beta N)^2} \equiv \frac{dy}{dt}, \quad (5.2)$$

defines a characteristic trajectory, along which

$$F = \frac{\alpha}{h}(N_e^2 - N^2) - \frac{2\beta N}{e\mu_0(N_e + \beta N)} \frac{db_i}{dy} \equiv \frac{dN}{dt}. \quad (5.3)$$

### 5.2.2 General Properties

The form of equation (5.1) means that structures in  $N$  are advected in the  $y$ -direction at a speed  $v_{IM}$ ; thus, the system supports waves that move at a speed  $v_{IM}$ . Since this is true of an ionosphere interacting with an ideal magnetosphere, we call this type of wave an ideal ionosphere-magnetosphere wave. Examining equation (5.2), we see that  $v_{IM}$  has the same sign as  $b_i$ ; noting the relationship between  $b_i$  and  $E_i$  (3.38), the corresponding velocity vector is

$$\begin{aligned} \mathbf{v}_{IM} &= v_{IM} \hat{\mathbf{y}} \\ &= \frac{2\beta N_e b_i}{e\mu_0 (N_e + \beta N)^2} \hat{\mathbf{y}} \\ &= \frac{2\beta N_e}{e\mu_0 v_A (N_e + \beta N)^2} E_i \hat{\mathbf{y}} \\ &= \frac{2\Sigma_{P0} N_e}{e (N_e + \beta N)^2} \mathbf{E}_i, \end{aligned} \quad (5.4)$$

so ionospheric density structures advect in the direction of the incident electric field (which is also the direction of the total electric field).

It is worth noting that  $v_{IM}$  depends on  $N$  such that regions of low ionospheric density are advected faster than regions of high ionospheric density. Therefore, troughs in  $N$  naturally catch up with crests. With an ideal magnetosphere, this leads to wavebreaking, collapsing the length-scale in  $y$  to zero and producing a current sheet in the magnetosphere. In practice, the length-scale will only collapse until electron inertia becomes significant in the magnetosphere; therefore, a complete theory should also consider electron inertial effects, which we treat in Section 5.3. The non-linear collapse of large length-scales to small length-scales is an important process, and we examine it in some detail in Section 5.4.

As density structures advect, they are subject to density changes, the structures obeying

$$\frac{dN}{dt} = F = \frac{\alpha}{h}(N_e^2 - N^2) - \frac{2\beta N}{e\mu_0(N_e + \beta N)} \frac{db_i}{dy},$$

where  $F$  is given by equation (5.3). Examining the terms on the right-hand side of the above

equation, these density changes are caused by the combined effects of ionisation, recombination and the rate of change of  $N$  due to field-aligned current that would occur if  $N$  were uniform.

In putting governing equation (3.46) into advection form (5.1), we have in fact separated the field-aligned current into two components, each of which plays a distinct role. Using Ampère's law (3.29) and the reflection coefficient (3.43,3.44) these components can be separated as,

$$\begin{aligned}
 j_z &= -\frac{1}{\mu_0} \frac{\partial b_x}{\partial y} \\
 &= -\frac{1}{\mu_0} \frac{\partial}{\partial y} ((1-r)b_i) \\
 &= \frac{b_i}{\mu_0} \frac{\partial r}{\partial y} - \frac{(1-r)}{\mu_0} \frac{db_i}{dy} \\
 &= \frac{b_i}{\mu_0} \frac{\partial}{\partial y} \left( \frac{N_e - \beta N}{N_e + \beta N} \right) - \frac{1}{\mu_0} \left( 1 - \frac{N_e - \beta N}{N_e + \beta N} \right) \frac{db_i}{dy} \\
 &= -\frac{2\beta N_e b_i}{\mu_0 (N_e + \beta N)^2} \frac{\partial N}{\partial y} - \frac{2\beta N}{\mu_0 (N_e + \beta N)} \frac{db_i}{dy}.
 \end{aligned} \tag{5.5}$$

The first term on the right-hand side causes the advection of density structures, and can be found in the characteristic velocity  $v_{IM}$  (5.2); it arises from gradients in  $N$  that in turn correspond to gradients in  $r$ . The second term alters the density structures as they advect, contributing to  $dN/dt$  on the characteristic trajectory (5.3); it is equal to the total (incident plus reflected) field-aligned current that would occur if  $N$  (and hence  $r$ ) were uniform. If a density structure is advected into a region where the incident Alfvén wave contains downward field-aligned current,  $N$  decreases as electrons are removed from the E-region to supply the magnetospheric current; conversely, if a density structure is advected into a region where the incident Alfvén wave contains upward field-aligned current,  $N$  increases as magnetospheric electrons are deposited into the E-region.

We can also investigate further the role played by ionisation and recombination. If  $N$  is perturbed from  $N_e$  so that  $N = N_e + \delta N$ , and we consider a uniform incident Alfvén wave that is current-free, then

$$\frac{d}{dt} (\delta N) = -\frac{\alpha}{h} \delta N (2N_e + \delta N), \tag{5.6}$$

$$\Rightarrow \delta N = \frac{(\delta N)_{t=0} \exp(-2\alpha N_e t/h)}{1 + ((\delta N)_{t=0}/2N_e)(1 - \exp(-2\alpha N_e t/h))}, \tag{5.7}$$

where the solution comes from separation of variables using partial fractions. If  $\delta N \ll 2N_e$  then this is simply exponential decay with an e-folding time

$$\tau_d = \frac{1}{2\alpha n_e}. \tag{5.8}$$

This same decay time can be obtained for ionospheric disturbances in the absence of active coupling to magnetosphere. Thus, the ionisation/recombination balance erodes structures in E-region

plasma-density as normal, regardless of whether or not these structures advect.

### 5.2.3 Advection and Damping of a Gaussian Wavepacket

The general properties of ideal IM-waves are readily confirmed using the numerical code described in Section 3.4. Here we will test two major predictions: (i) IM-waves advect in the direction of the electric field; and (ii) perturbations are damped by ionisation and recombination. The code was tailored to this study using periodic boundary-conditions in  $\tilde{y}$ , a uniform incident Alfvén wave ( $\tilde{u}_i = -1$  everywhere, corresponding to negative  $\tilde{E}_y$ ), and a small-amplitude disturbance in  $\tilde{N}$  to eliminate non-linear effects.

A suitable choice for the initial density perturbation is a Gaussian wavepacket:

$$\tilde{N} = 1 + \delta\tilde{N}_{amp} \exp\left(-\frac{(\tilde{y} - \tilde{y}_c)^2}{2\sigma^2}\right). \quad (5.9)$$

This waveform is initially centred at  $\tilde{y}_c$ , has amplitude  $\delta\tilde{N}_{amp}$ , and has a full-width at half maximum (FWHM) given by

$$\text{FWHM} = 2\sqrt{2\ln(2)}\sigma \approx 2.35\sigma. \quad (5.10)$$

Images of the wavepacket, positioned outside the domain, are included in the initial condition to satisfy periodic boundary-conditions at  $\tilde{t} = 0$ . Our domain extends from  $\tilde{y} = -1$  to  $\tilde{y} = 1$ , so a suitable number of images is included by setting

$$\tilde{N} = 1 + \delta\tilde{N}_{amp} \sum_{j=-g}^g \exp\left(-\frac{(\tilde{y} - \tilde{y}_c - 2j)^2}{2\sigma^2}\right). \quad (5.11)$$

For the simulations presented here, parameters were set as  $\tilde{y}_c = 0.8$ ,  $\sigma = 0.05 \Rightarrow \text{FWHM} \approx 0.118$ ,  $\delta\tilde{N}_{amp} = 0.001$  and  $g = 1$  to include one image on either side of the wavepacket.

In order to simultaneously demonstrate advection and damping, we must ensure that both processes occur on reasonable time-scales within the simulation. Applying the normalisations of Section 3.4.1 to the advection speed (5.2), putting  $b_i = B_0 u_i / v_A = -B_0 u_{i0} / v_A$  for a uniform incident Alfvén wave, and using the definition of  $\eta$  (3.54), one can show that

$$\tilde{v}_{IM} = -\frac{\eta}{(1 + \beta\tilde{N})^2}. \quad (5.12)$$

In these simulations it is convenient to have  $\beta = 10$  and  $\tilde{v}_{IM} = -0.1$  where  $\tilde{N} = 1$ ; we therefore set  $\eta = 12.1$ .

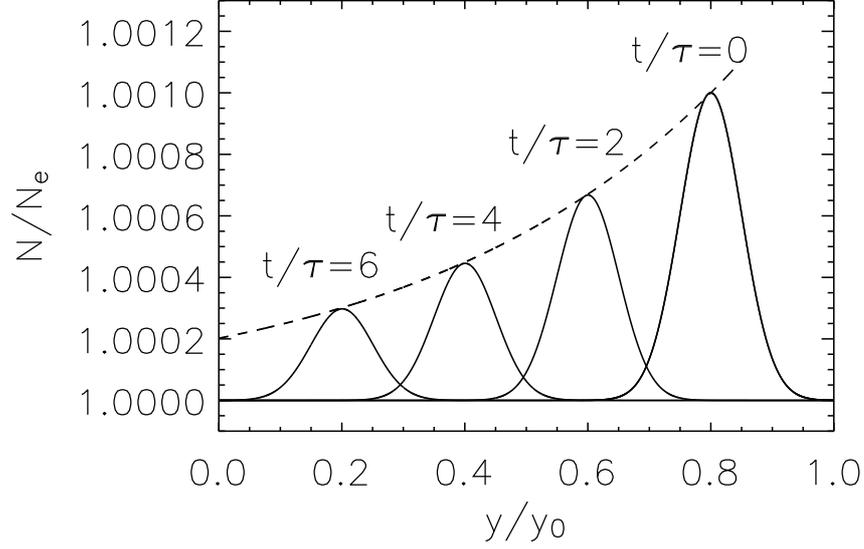


Figure 5.1: Advection of a Gaussian wavepacket under an ideal magnetosphere. The dashed curve plots exponential decay with an e-folding time  $\tau_d = 1/(2\alpha n_e)$ .

Normalising the decay time, equation (5.8) becomes

$$\tilde{\tau}_d = \frac{1}{2\tilde{\alpha}}. \quad (5.13)$$

A choice of  $\tilde{\alpha} = 0.1 \Rightarrow \tilde{\tau}_d = 5$  is a good choice that ensures both advection and damping are clearly seen in simulation results.

Results are shown in Figure 5.1, which plots the solution at four different times:  $\tilde{t} = 0, 2, 4$  and  $6$ . During the simulation, the wavepacket moves from its initial position, centred on  $\tilde{y} = 0.8$ , to a new position  $\tilde{y} = 0.2$ , and the original shape of the wavepacket is preserved. This motion takes place in the direction of the incident electric field and corresponds to a speed  $\tilde{v} = -0.1$ .

During advection, the amplitude of the wavepacket decays significantly. Theory predicts that the wave damps with an e-folding time given by (5.13), and moves at a speed  $\tilde{v} = -0.1$ . By combining these assumptions, it is possible to predict the amplitude of the wavepacket at any given location: the appropriate curve is plotted as a dashed line on Figure 5.1 and it reveals excellent agreement between theory and simulation.

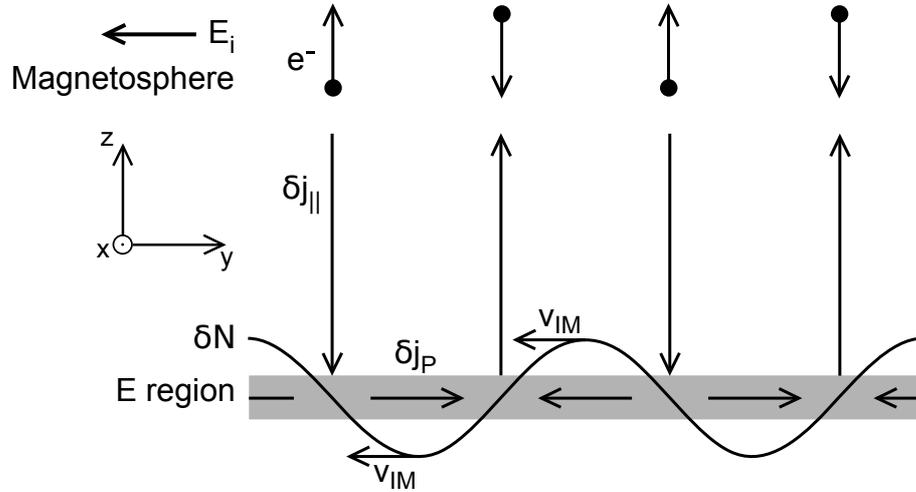


Figure 5.2: Cartoon of an ionosphere-magnetosphere wave. Gradients in height-integrated ionospheric number density produce field-aligned currents in the magnetosphere by reflection of an incident Alfvén wave. These, in turn, modify the ionospheric number density, leading to the advection of wave forms.

#### 5.2.4 Physical Cartoon

The physics of advection can be illustrated with a simple cartoon, as shown in Figure 5.2. For simplicity, we consider a uniform Alfvén wave incident on a small amplitude, sinusoidal disturbance in  $N$ . Recombination and ionisation are neglected, and an incident electric field is assumed, directed to the left.

The key principle behind advection is that gradients in  $N$  produce field-aligned currents (5.5), which subsequently modify  $N$  (3.2). In the cartoon example, positive gradients in  $N$  lead to the deposition of more electrons, and negative gradients in  $N$  lead to the removal of electrons. Meanwhile, Pedersen current moves E-region ions from negative gradients in  $N$  to positive gradients. In this way, positive gradients of  $N$  become the new peaks and negative gradients become the new troughs, so the waveform advects.

### 5.3 Inertial IM-Waves

#### 5.3.1 Normal-Mode Analysis

When electron inertia (i.e. non-zero electron mass) is considered for the magnetosphere, Alfvén waves become dispersive at small length-scales, and an advection equation equivalent to (5.1) no longer exists. In this case, the most revealing treatment is a linear normal-mode analysis.

Consider a steady-state E-region plasma-density  $N_{SS}$ , corresponding to some incident Alfvén wave, and perturb  $N$  so that

$$N = N_{SS} + \delta N \quad (5.14)$$

where  $\delta N$  is a small perturbation. At the top of the E-region, reflection of the incident Alfvén wave from the E-region produces total (incident plus reflected) fields

$$b_x = b_{x0} + \delta b_x, \quad (5.15)$$

$$E_y = E_{y0} + \delta E_y, \quad (5.16)$$

where  $b_{x0}$  and  $E_{y0}$  are the total fields that exist in the steady-state.

We wish to solve equations (3.2) and (3.6) in tandem. Repeating these here, (3.2) can be written as

$$\frac{\partial N}{\partial t} + \frac{1}{e\mu_0} \frac{\partial b_x}{\partial y} = \frac{\alpha}{h} (N_e^2 - N^2),$$

and (3.6) is

$$b_x = \frac{\mu_0 \Sigma_{P0}}{N_e} N E_y.$$

Note that in the steady state, these become

$$\frac{1}{e\mu_0} \frac{\partial b_{x0}}{\partial y} = \frac{\alpha}{h} (N_e^2 - N_{SS}^2), \quad (5.17)$$

$$b_{x0} = \frac{\mu_0 \Sigma_{P0}}{N_e} N_{SS} E_{y0}; \quad (5.18)$$

both results being of later use.

Equations (5.14) and (5.16) can be used to rewrite (3.2), yielding

$$\begin{aligned} \frac{\partial}{\partial t} (N_{SS} + \delta N) + \frac{1}{e\mu_0} \frac{\partial}{\partial y} (b_{x0} + \delta b_x) &= \frac{\alpha}{h} (N_e^2 - (N_{SS} + \delta N)^2) \\ \Rightarrow \frac{\partial}{\partial t} (\delta N) + \frac{1}{e\mu_0} \frac{\partial b_{x0}}{\partial y} + \frac{1}{e\mu_0} \frac{\partial}{\partial y} (\delta b_x) &= \frac{\alpha}{h} (N_e^2 - N_{SS}^2) - \frac{2\alpha N_{SS}}{h} \delta N - \frac{2\alpha}{h} \delta N^2 \\ &\Rightarrow \frac{\partial}{\partial t} (\delta N) + \frac{1}{e\mu_0} \frac{\partial}{\partial y} (\delta b_x) = -\frac{2\alpha N_{SS}}{h} \delta N - \frac{2\alpha}{h} \delta N^2. \end{aligned} \quad (5.19)$$

Similarly, (3.6) can be rewritten as

$$\begin{aligned} (b_{x0} + \delta b_x) &= \frac{\mu_0 \Sigma_{P0}}{N_e} (N_{SS} + \delta N) (E_{y0} + \delta E_y) \\ \Rightarrow b_{x0} \left(1 + \frac{\delta b_x}{b_{x0}}\right) &= \frac{\mu_0 \Sigma_{P0} N_{SS} E_{y0}}{N_e} \left(1 + \frac{\delta N}{N_{SS}}\right) \left(1 + \frac{\delta E_y}{E_{y0}}\right) \end{aligned}$$

$$\begin{aligned}
\Rightarrow 1 + \frac{\delta b_x}{b_{x0}} &= \left(1 + \frac{\delta N}{N_{SS}}\right) \left(1 + \frac{\delta E_y}{E_{y0}}\right) \\
\Rightarrow \frac{\delta b_x}{b_{x0}} &= \frac{\delta N}{N_{SS}} + \frac{\delta E_y}{E_{y0}} + \frac{\delta N \delta E_y}{N_{SS} E_{y0}}.
\end{aligned} \tag{5.20}$$

Progress can be made if we seek linear normal-modes proportional to  $\exp(i(\mathbf{k} \cdot \mathbf{r} - \omega t))$ , so that products of perturbations are negligible, and derivatives of perturbations become  $\partial/\partial y \equiv ik_y$ ,  $\partial/\partial z \equiv ik_z$  and  $\partial/\partial t \equiv -i\omega$  (the angular frequency,  $\omega$ , may be complex, but we assume  $k_y$  to be real). This assumption can be made rigorously under the condition that  $N_{SS}$  is constant across the whole domain. Since we require field-aligned current to be zero at the edges of our domain, there is only one constant steady-state:  $N_{SS} = N_e$ . I therefore ask the reader to bear in mind that, strictly speaking, the following analysis is only rigorous if the incident Alfvén wave is current-free.

As we shall see in Chapter 6, inertial IM-waves naturally occur in large-scale systems of field-aligned current, and we would like to describe their properties in this case too. It is possible to extend the normal-mode analysis to cover this, if we allow it to be a local analysis; that is to say, we assume normal-modes that vary rapidly in  $y$ , so that the variation in the steady-state is negligible by comparison. Due to the power of local analysis to cope with a wide range of scenarios (including  $N_{SS} = N_e$ ), we proceed on these more general lines. Although approximate, this approach is very successful in describing the results of Chapter 6.

Under the assumption of linear normal-modes, perturbations  $\delta b_x$  and  $\delta E_y$  form an upgoing inertial Alfvén wave and satisfy

$$\delta E_y = -v_A \sqrt{1 + k_y^2 \lambda_e^2} \delta b_x, \tag{5.21}$$

where

$$\lambda_e = \sqrt{\frac{m^- m^+}{\mu_0 \rho_0 e^2}} \equiv \sqrt{\frac{m^-}{\mu_0 n_m e^2}} \tag{5.22}$$

is the electron inertial length at the base of the magnetosphere.

A relation between  $\delta b_x$  and  $\delta N$  can be found if we linearise (5.20), then eliminate  $\delta E_y$  using (5.21) and  $E_{y0}$  using (5.18):

$$\frac{\delta b_x}{b_{x0}} = \frac{\delta N}{N_{SS}} + \frac{\delta E_y}{E_{y0}} \tag{5.23}$$

$$\Rightarrow \frac{\delta b_x}{b_{x0}} = \frac{\delta N}{N_{SS}} - \left(v_A \sqrt{1 + k_y^2 \lambda_e^2} \delta b_x\right) \left(\frac{\mu_0 \Sigma_{P0} N_{SS}}{N_e b_{x0}}\right) \tag{5.24}$$

$$\Rightarrow \frac{\delta b_x}{b_{x0}} = \frac{\delta N}{N_{SS}} - \frac{\beta N_{SS}}{N_e} \sqrt{1 + k_y^2 \lambda_e^2} \frac{\delta b_x}{b_{x0}} \tag{5.25}$$

$$\Rightarrow \frac{\delta b_x}{b_{x0}} \left( 1 + \frac{\beta N_{SS}}{N_e} \sqrt{1 + k_y^2 \lambda_e^2} \right) = \frac{\delta N}{N_{SS}} \quad (5.26)$$

$$\Rightarrow \delta b_x = \frac{N_e b_{x0} \delta N}{N_{SS} \left( N_e + \beta N_{SS} \sqrt{1 + k_y^2 \lambda_e^2} \right)}. \quad (5.27)$$

Using this to remove  $\delta b_x$  from equation (5.19), and dropping non-linear terms yields

$$\frac{\partial}{\partial t} (\delta N) + \frac{1}{e\mu_0} \frac{\partial}{\partial y} \left( \frac{N_e b_{x0} \delta N}{N_{SS} \left( N_e + \beta N_{SS} \sqrt{1 + k_y^2 \lambda_e^2} \right)} \right) = -\frac{2\alpha N_{SS}}{h} \delta N. \quad (5.28)$$

Assuming that  $\delta N$  varies much more rapidly in  $y$  than the background (steady-state) quantities, background quantities can be taken outside the  $y$ -derivative, giving

$$\frac{\partial}{\partial t} (\delta N) + \frac{N_e b_{x0}}{e\mu_0 N_{SS} \left( N_e + \beta N_{SS} \sqrt{1 + k_y^2 \lambda_e^2} \right)} \frac{\partial}{\partial y} (\delta N) = -\frac{2\alpha N_{SS}}{h} \delta N. \quad (5.29)$$

Next, since  $\delta N$  is a normal-mode, derivatives are replaced as previously discussed, giving

$$-i\omega \delta N + \frac{N_e b_{x0}}{e\mu_0 N_{SS} \left( N_e + \beta N_{SS} \sqrt{1 + k_y^2 \lambda_e^2} \right)} i k_y \delta N = -\frac{2\alpha N_{SS}}{h} \delta N,$$

and some rearranging yields the dispersion relation

$$\omega = \frac{b_{x0}}{e\mu_0 N_{SS} (\beta N_{SS}/N_e) \lambda_e} \frac{(\beta N_{SS}/N_e) k_y \lambda_e}{\left( 1 + (\beta N_{SS}/N_e) \sqrt{1 + k_y^2 \lambda_e^2} \right)} - i \frac{2\alpha N_{SS}}{h}. \quad (5.30)$$

### 5.3.2 General Properties and Dispersion Diagram

The dispersion relation (5.30) provides a lot of information about IM-waves. Since  $\omega$  is complex, we will examine the real and imaginary parts separately.

The imaginary part of (5.30) reveals growth or damping. Here perturbations decay exponentially with an e-folding time

$$\tau_d = \frac{1}{2\alpha n_{SS}}. \quad (5.31)$$

This suggests that damping of perturbations by ionization/recombination in the E-region is unaffected by electron-inertia in the magnetosphere.

The real part of (5.30) shows the oscillatory properties of the solution. We take the real part as

$$\omega_r = \omega_{IM} \frac{(\beta N_{SS}/N_e) k_y \lambda_e}{\left(1 + (\beta N_{SS}/N_e) \sqrt{1 + k_y^2 \lambda_e^2}\right)}, \quad (5.32)$$

where

$$\omega_{IM} = \frac{b_{x0}}{e\mu_0 N_{SS} (\beta N_{SS}/N_e) \lambda_e}. \quad (5.33)$$

Note that the term  $\beta N_{SS}/N_e$  is the ratio of Pedersen conductance to ideal Alfvén conductance in the steady state, since

$$\beta \frac{N_{SS}}{N_e} = \frac{\Sigma_{P0} N_{SS}}{\Sigma_A N_e} = \frac{1}{\Sigma_A} \frac{\Sigma_{P0} N_{SS}}{N_e} = \frac{\Sigma_{P,SS}}{\Sigma_A}. \quad (5.34)$$

The general behaviour of IM-waves, subject to electron inertial effects in the magnetosphere, is revealed by a plot of  $\omega_r/\omega_{IM}$  against  $k_y \lambda_e$  (the real part of the dispersion diagram). This is provided in Figure 5.3. Two speeds are important in interpreting the dispersion diagram. At any point on the curve, the gradient of a straight line passing through that point and the origin is  $v_{ph} \omega_{IM}/\lambda_e$  where  $v_{ph} = \omega_r/k_y$  is the phase-speed. The slope of the curve is  $v_g \omega_{IM}/\lambda_e$ , where  $v_g = \partial\omega_r/\partial k_y$  is the group-speed of the waves.

In the limit of large length-scales ( $k_y \lambda_e \rightarrow 0$ ), group and phase-speeds are equal to one another and independent of  $k_y$ . In this limit

$$\frac{\partial\omega_r}{\partial k_y} = \frac{\omega_r}{k_y} = \frac{b_{x0}}{e\mu_0 N_{SS} (1 + \beta N_{SS}/N_e)}. \quad (5.35)$$

Since  $b_{x0} = (1 - r_{SS})b_i$ , it is equivalent to write

$$\begin{aligned} v_{ph,g} &\rightarrow \frac{2(\beta N_{SS}/N_e) b_i}{e\mu_0 N_{SS} (1 + \beta N_{SS}/N_e)^2} \\ \Rightarrow v_{ph,g} &\rightarrow \frac{2\beta N_e b_i}{e\mu_0 (N_e + \beta N_{SS})^2} \\ \Rightarrow v_{ph,g} &\rightarrow v_{IM,lin}, \end{aligned} \quad (5.36)$$

where  $v_{IM,lin}$  is the linear limit of equation (5.2) obtained by putting  $N \rightarrow N_{SS}$ . It follows that structures in E-region plasma-density with  $k_y \lambda_e \ll 1$  advect in  $y$  at a speed  $v_{IM,lin}$ ; hence the normal-mode analysis recovers the results of Section 5.2 for ideal, linear disturbances.

The character of IM-waves changes substantially if we consider strongly inertial length-scales ( $k_y \lambda_e \gtrsim 2\pi$ ). In this limit, the group-speed goes to zero: disturbances remain in a fixed location

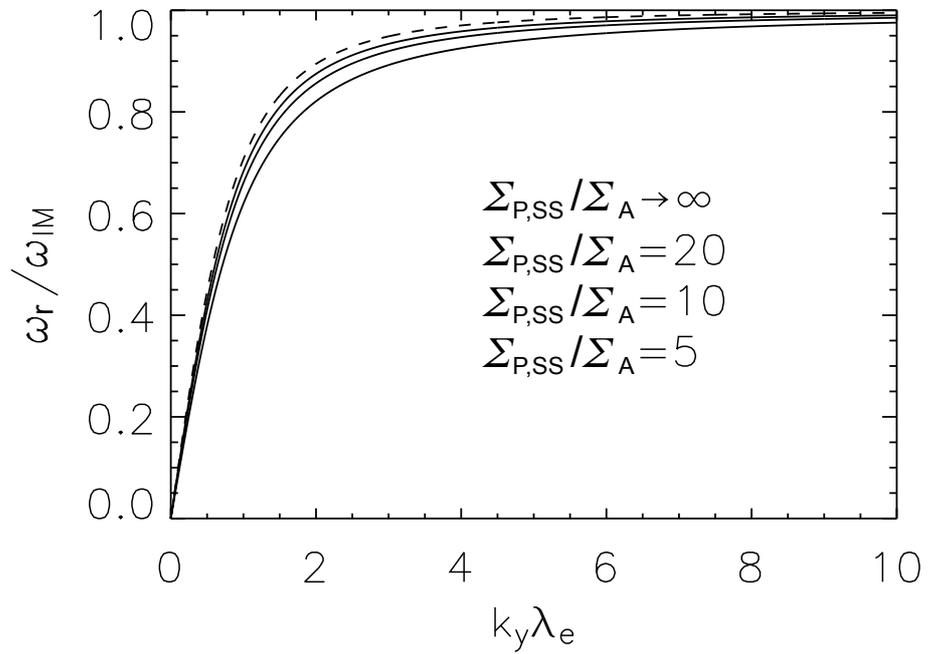


Figure 5.3: Dispersion diagram for linear normal-mode analysis including the effects of electron inertia in the magnetosphere. Dashed curve:  $\Sigma_{P,SS}/\Sigma_A \rightarrow \infty$ . Solid curves:  $\Sigma_{P,SS}/\Sigma_A = 20, 10$  and  $5$ , frequency increasing with  $\Sigma_{P,SS}/\Sigma_A$ . For large length-scales (small  $k_y\lambda_e$ ),  $\omega/k_y = v_{IM}$  is independent of  $k_y$  so density disturbances are advected at this speed. For small length-scales ( $k_y\lambda_e \gtrsim 2\pi$ ), density disturbances oscillate around  $N = N_e$  with a period  $\tau_{IM} = 2\pi/\omega_{IM}$ .

and oscillate about  $N = N_{SS}$  with a period that approaches

$$\tau_{IM} = \frac{2\pi}{\omega_{IM}} = \frac{2\pi\epsilon\mu_0 N_{SS}(\beta N_{SS}/N_e)\lambda_e}{b_{x0}} \quad (5.37)$$

from above. The phase-speed typically remains non-zero, so we expect to see a phase-motion in the waves. (Phase-motion gives the appearance that density structures are moving, as different parts of the disturbance oscillate with different phase, however phase-motion cannot carry the disturbance into previously undisturbed regions.) A simulation of IM-waves in the strongly inertial limit is presented in the next section, 5.3.3.

### 5.3.3 Evolution of a Strongly Inertial Wavepacket

It is useful to visualise IM-waves with computer simulations, as we have already done for the ideal limit; now let us investigate a wavepacket in the strongly inertial regime, using the computer code described in Section 3.5.

For the study of IM-waves, boundaries in  $\tilde{y}$  were made periodic, using ghost-cells. If the domain contains  $ny$  physical grid-points in the  $y$ -direction, then these are assigned indices 2 to  $ny + 1$ , with ghosts at grid-points 1 and  $ny + 2$ . Periodic boundaries are achieved by copying the physical value of 3 to the ghost at  $ny + 2$ , and the physical value of  $ny - 1$  to the ghost at 1. This way, the physical grid-points at 2 and  $ny$  become clones of one another.

The system was driven using an incident Alfvén wave that is uniform in  $\tilde{y}$  and initially ramped between  $\tilde{z} = 0$  and  $\tilde{z} = 0.3$ . There is therefore a short transient period between  $\tilde{t} = 0$  and  $\tilde{t} = 0.3$  during which the incident Alfvén wave at the top of the E-region increases to its maximum amplitude. At full amplitude, the incident magnetic field perturbation is  $\tilde{b}_i = -1$ , corresponding to an electric field perturbation  $\tilde{E}_i = -1$ .

Lastly, the initial condition for  $\tilde{N}$  was set as a Gaussian wavepacket, using the techniques outlined in Section 5.2.3. Since the physical domain extends from  $\tilde{y} = 0$  to  $\tilde{y} = 1$ , E-region plasma-density was initialised as

$$\tilde{N} = 1 + \delta\tilde{N}_{amp} \sum_{j=-1}^1 \exp\left(-\frac{(\tilde{y} - \tilde{y}_c - j)^2}{2\sigma^2}\right), \quad (5.38)$$

which includes two images to satisfy periodic boundary-conditions. For the simulation presented here,  $\sigma = 0.05 \Rightarrow \text{FWHM} \approx 0.118$ .

Theory predicts that strongly inertial IM-waves ( $k_y\lambda_e \gg 2\pi$ ) behave very differently to their ideal counterparts, remaining stationary (since  $v_g \approx 0$ ) and oscillating at a period approaching  $\tau_{IM}$  from above. This regime is accessed in the simulation by setting  $\tilde{\lambda}_e = 0.5$ , so that the electron

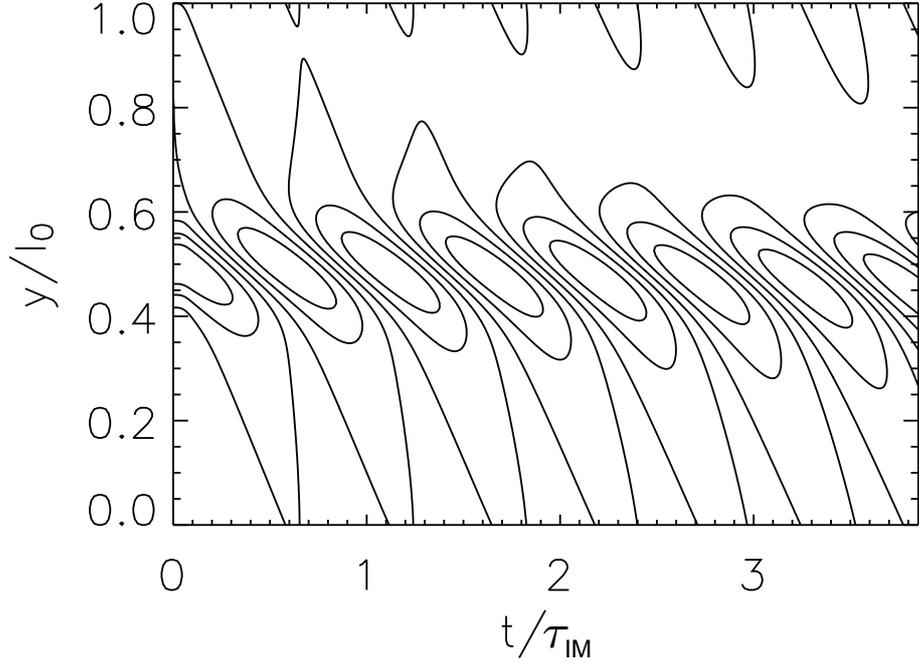


Figure 5.4: Contour plot of  $\tilde{N}$ , showing evolution of a wavepacket in the strongly inertial regime ( $\tilde{\lambda}_e = 0.5$ ) without recombination. The dominant effect is an oscillation with period approaching  $|\tau_{IM}|$ .

inertial length is more than four times the FWHM of the initial wavepacket. We are not expecting the wavepacket to move substantially from its initial position, so it is centred in the domain at  $\tilde{y}_c = 0.5$ . Finally, because we are testing a linear theory, the amplitude of the wavepacket was set as  $\tilde{N}_{amp} = 0.001$ , and we turn off ionisation and recombination to exclude damping.

Results are shown in Figure 5.4 as a contour plot of  $\tilde{N}$ . The wavepacket behaves as expected, remaining relatively stationary, and oscillating at a characteristic frequency. We expect the period of oscillation to be slightly longer than  $|\tau_{IM}| = 2\pi/|\omega_{IM}|$  where  $\omega_{IM}$  is given by equation (5.33); applying the normalisations of Section 3.5, one can show that

$$|\tilde{\omega}_{IM}| = \frac{\epsilon\eta}{\beta(1+\beta)\tilde{\lambda}_e} \quad (5.39)$$

$$\Rightarrow |\tilde{\tau}_{IM}| = \frac{2\pi\beta(1+\beta)\tilde{\lambda}_e}{\epsilon\eta}, \quad (5.40)$$

which gives  $|\tilde{\tau}_{IM}| = 10\pi/11 \approx 2.86$  for  $\beta = 10$ ,  $\tilde{\lambda}_e = 0.5$  and  $\epsilon\eta = 121$ . For easy comparison with this prediction, the time axis of Figure 5.4 is plotted as  $t/|\tau_{IM}| \equiv \tilde{t}/|\tilde{\tau}_{IM}|$ . Examining the contour plot, each cycle of oscillation lasts approximately  $T = 1.1\tau_{IM}$ , which agrees very well with the theory of Section 5.3.1.

Now let us consider velocities. Taking group speed first, a small drift is apparent in Figure

5.4 if one looks over long enough time-scales: this drift follows the direction of the electric field, and (for the peak of the wavepacket) amounts to a distance  $\tilde{D} \approx 0.05$  over three full cycles of oscillation. Note also that a Gaussian wavepacket contains a range of different spatial scales, and the longest wavelength components (smallest  $k_y \lambda_e$ ) are seen at the edge of the waveform, which drifts more rapidly than the centre of the waveform. All of these features are ‘predicted’ by the dispersion diagram shown in Figure 5.3, which shows group speed going (almost but not quite) to zero in the strongly inertial regime.

We can also consider phase-speed, apparent as ‘diagonal stripes’ in Figure 5.4. Intuitively, one might have expected the wavepacket to oscillate in phase, but this is not the case. The different phases of oscillation produce a phase-velocity in the direction of the electric field, in good qualitative agreement with the theory.

## 5.4 Multi-Scale Coupling by Wavebreaking of IM-waves

In Section 5.2, we explored the advective nature of ideal (large-scale) IM-waves. The speed of advection (5.2) depends on the height-integrated number density  $N$ , such that troughs in  $N$  naturally advect faster than peaks in  $N$ . The difference is negligible for linear IM-waves ( $\delta N \ll N_e$ ), so these advect without changing their original shape; however, the shearing effect of different advection speeds can be considerable for non-linear waves. It follows that large-amplitude, large-scale ( $\lambda_y \gg \lambda_e$ ) enhancements in  $N$  will steepen their trailing-edge as they advect. With an ideal magnetosphere, the waves steepen to a discontinuity; if electron inertial effects are included in the magnetosphere, then the steepening becomes a source of inertial IM-waves as the width of the steepened section approaches the electron inertial length.

### 5.4.1 Wavebreaking of Ideal IM-waves

#### 5.4.1.1 Numerical Simulation

Our investigation of wavebreaking begins with a numerical simulation showing the evolution of an ideal non-linear wavepacket. The code is identical to that used in Section 5.2.3, but we change two parameters: first, we set  $\tilde{\alpha} = 0$  to exclude damping by recombination and ionisation; and second, we increase the amplitude of the initial waveform.

The amplitude of the initial wavepacket is chosen so that wavebreaking occurs on a reasonable simulation time-scale, compared with advection. Examining equation (5.12), if background density ( $\tilde{N} = 1$ ) corresponds to an advection speed  $\tilde{v}_0$ , then the peak of the Gaussian packet moves

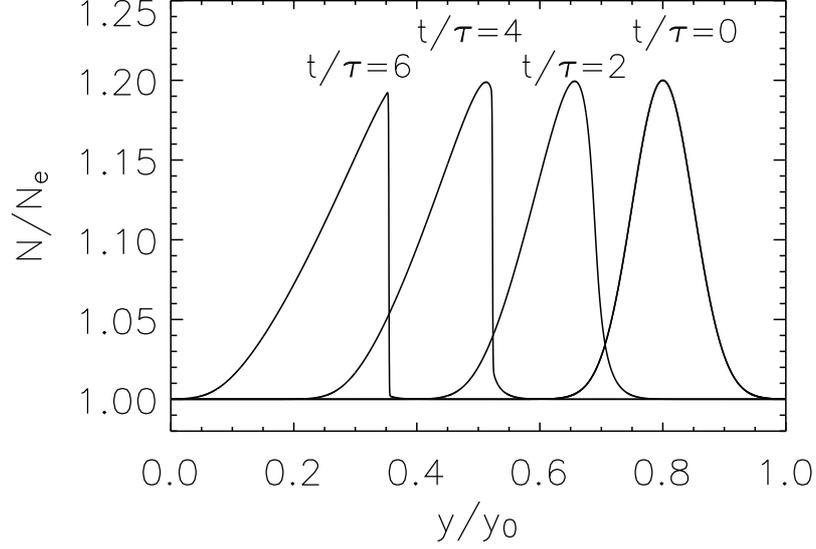


Figure 5.5: Wavebreaking of a large-amplitude Gaussian wavepacket assuming an ideal magnetosphere. Since troughs in height integrated number density advect faster than crests, they ‘catch up’ on the trailing-edge of the peak and the wave breaks. Here, damping by ionisation and recombination has been artificially switched off by setting  $\tilde{\alpha} = 0$ .

at

$$\tilde{v}_{peak} = -\frac{\eta}{\left(1 + \beta \left(1 + \delta\tilde{N}_{amp}\right)\right)^2} = \frac{\tilde{v}_0(1 + \beta)^2}{\left(1 + \beta \left(1 + \delta\tilde{N}_{amp}\right)\right)^2} \quad (5.41)$$

Meanwhile, the Gaussian perturbation drops from its peak to one percent of its peak over a distance

$$\tilde{D} = \sqrt{2 \ln(100)} \sigma. \quad (5.42)$$

It follows that background density to the right of the wavepacket should catch up with the peak over a time-scale

$$\tilde{\tau}_{catch-up} = \frac{\tilde{D}}{(\tilde{v}_{peak} - \tilde{v}_0)} \approx 5.35, \quad (5.43)$$

for  $\sigma = 0.05$ ,  $\tilde{v}_0 = -0.1$  and  $\delta\tilde{N}_{peak} = 0.2$ . Thus,  $\delta\tilde{N}_{peak} = 0.2$  is a good choice for the initial amplitude of the wavepacket, since it produces wavebreaking on a useful time-scale.

Figure 5.5 shows results from the simulation at  $\tilde{t} = 0, 2, 4$  and  $6$ . The packet advects to the left, as in the small-amplitude case (Figure 5.1), but now the peak travels noticeably slower than the background, so that the wavepacket becomes distorted. Wavebreaking is inevitable, and this

occurs at the trailing-edge of the wavepacket between  $\tilde{t} = 3$  and  $\tilde{t} = 4$ . By  $\tilde{t} = 6$ , the background density has caught up with the peak of the disturbance, as predicted by theory.

#### 5.4.1.2 Speed of Discontinuity

We can obtain a formula for the speed of the discontinuity by integrating equation (3.46) over a small interval in  $y$  around the discontinuity, in the frame of the discontinuity. In the ‘lab’ frame, (3.46) states

$$\frac{\partial N}{\partial t} + \frac{1}{\mu_0 e} \frac{\partial}{\partial y} \left( \frac{2\beta N b_i}{N_e + \beta N} \right) = \frac{\alpha}{h} (N_e^2 - N^2).$$

For a discontinuity that moves at a speed  $U(t)$  in the lab frame, the two frames are related by

$$y' = y - Ut, \quad (5.44)$$

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + U \frac{\partial}{\partial y}, \quad (5.45)$$

$$\frac{\partial}{\partial y'} = \frac{\partial}{\partial y}, \quad (5.46)$$

where primes denote variables in the frame of the discontinuity. Thus, in the frame of the discontinuity,

$$\begin{aligned} \left( \frac{\partial}{\partial t'} - U \frac{\partial}{\partial y'} \right) N + \frac{1}{\mu_0 e} \frac{\partial}{\partial y'} \left( \frac{2\beta N b_i}{N_e + \beta N} \right) &= \frac{\alpha}{h} (N_e^2 - N^2) \\ \Rightarrow \frac{\partial}{\partial y'} \left( -UN + \frac{2\beta N b_i}{\mu_0 e (N_e + \beta N)} \right) &= \frac{\alpha}{h} (N_e^2 - N^2) - \frac{\partial N}{\partial t'}. \end{aligned} \quad (5.47)$$

To obtain  $U$  for the discontinuity, we integrate (5.47) over a small region of width  $2\epsilon$ , centred the location of the discontinuity,  $\xi'$ . Doing so, and taking the limit  $\epsilon \rightarrow 0$ ,

$$\begin{aligned} \int_{\xi' - \epsilon}^{\xi' + \epsilon} \frac{\partial}{\partial y'} \left( -UN + \frac{2\beta N b_i}{\mu_0 e (N_e + \beta N)} \right) dy' &= \int_{\xi' - \epsilon}^{\xi' + \epsilon} \left( \frac{\alpha}{h} (N_e^2 - N^2) - \frac{\partial N}{\partial t'} \right) dy' \\ \Rightarrow \left[ -UN + \frac{2\beta N b_i}{\mu_0 e (N_e + \beta N)} \right]_{-}^{+} &= \text{terms of order } \epsilon \rightarrow 0, \end{aligned} \quad (5.48)$$

where the  $+$  ( $-$ ) on the square bracket denotes that quantities are to be evaluated to the right (left) of the discontinuity. If we introduce

$$N^+ = \lim_{y \rightarrow \xi^+} \{N\}, \quad (5.49)$$

$$N^- = \lim_{y \rightarrow \xi^-} \{N\}, \quad (5.50)$$

then because  $N$  is the only non-continuous quantity in (5.48), rearranging for  $U$  yields,

$$U = \frac{2\beta N_e b_i}{\mu_0 e (N_e + \beta N^+) (N_e + \beta N^-)}. \quad (5.51)$$

Comparing this to the advection speed given by (5.2), the speed of a discontinuity can be expressed as

$$U = \text{sign}(b_i) \sqrt{v_{IM}^+ v_{IM}^-} = \text{sign}(E_i) \sqrt{v_{IM}^+ v_{IM}^-}. \quad (5.52)$$

Thus, once formed, a discontinuity travels in the direction of the electric field, at a speed given by the geometric mean of advection speeds on either side of the discontinuity.

The expression (5.51) is easily checked against the numerical simulation. To do so, we identify the discontinuity as the greatest difference in  $N$  between two adjacent grid-points, making sure that this corresponds to a negative gradient in  $N$  (the discontinuity forms at a negative gradient because advection occurs from right to left). Tracking this location in time gives the dimensionless speed of the discontinuity, and also the values of  $\tilde{N}^+$ ,  $\tilde{N}^-$  and  $\tilde{b}_i(\tilde{y})$ . The tracked speed and the speed obtained from (5.51) agree very well, the difference between the two typically being within 0.5% of their mean, which appears consistent with the uncertainty inherent from tracking discrete data.

#### 5.4.2 Wavebreaking as a Source of Inertial IM-waves

When a large-scale IM-wave steepens, it provides a direct route to the small length-scales at which electron inertia becomes important. A simulation readily demonstrates this multi-scale evolution, requiring just a few changes to the computer code used in Section 5.3.3. Here we reuse the initial condition from the ideal wavebreaking simulation of Section 5.4.1, with  $\tilde{N}_{amp} = 0.2$ ,  $\tilde{y}_c = 0.8$  and  $\sigma = 0.05$  as before. Parameters are set to give  $\tilde{v}_0 = -0.1$  with  $\beta = 10$  (following the ideal simulation) and the electron inertial length,  $\tilde{\lambda}_e = 0.01$ , is set small with the aim of seeing the wavepacket advect and steepen at its trailing-edge, before electron inertia introduces dispersion.

The results are summarised by Figure 5.6, which shows snapshots of  $\tilde{N}$  at five different times separated by a fixed interval. Initially, the wavepacket advects as in the ideal case, steepening at its trailing-edge as the background density catches up with the peak. In time, the trailing-edge steepens enough that electron inertial effects become significant, and a series of undershoots and overshoots develops behind the wavepacket.

It is clear that the undershoot immediately behind the wavepacket grows in time, and its growth is plotted in Figure 5.7. Examining Figure 5.7, the undershoot begins to grow at about  $\tilde{t} = 1.5$ , and we can identify this as the time at which steepening of the trailing edge first accesses iner-

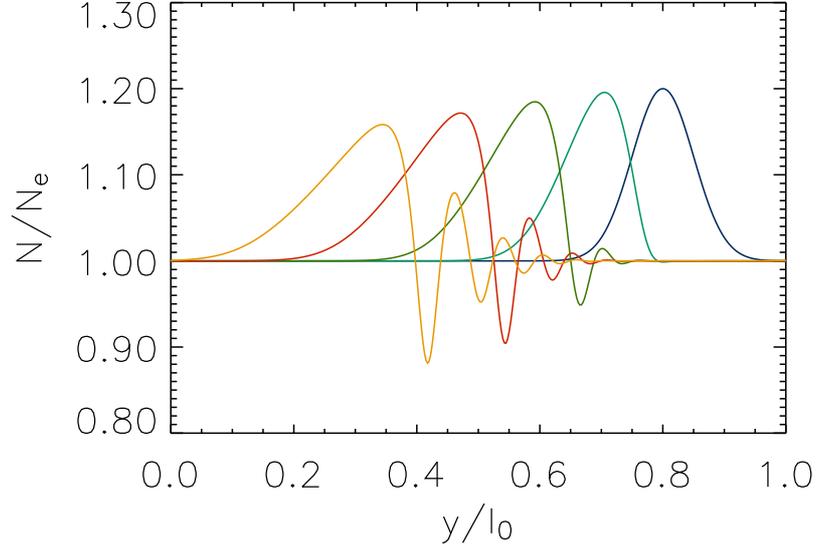


Figure 5.6: Snapshots showing the evolution of a large-scale non-linear wavepacket subject to electron inertial effects in the magnetosphere. Snapshots are shown at  $\tilde{t} = 0$  (dark blue), 1.5 (turquoise), 3 (green), 4.5 (red) and 6 (gold).

tial length-scales. Subsequent growth of the undershoot is rapid as non-linear advection transfers power from the large wavelengths of the main wavepacket to the small wavelengths of the trailing oscillatory structure. Growth, however, can only continue as long as there is a large-scale wavepacket to feed into the trailing-edge: at later times, as the main wavepacket is consumed, growth of the undershoot slows.

Taking a closer look at the small-scale waves, Figure 5.8 shows a snapshot of  $\tilde{N}$  at  $\tilde{t} = 6$ , with a range chosen to focus on these waves. Two vertical dashed lines are also plotted, separated by  $2\pi\tilde{\lambda}_e$ , which we can see as the approximate wavelength, implying  $k_y\lambda_e \approx 1$ . Referring to the dispersion diagram (Figure 5.3), this places the waves in a transition-region in  $k_y\lambda_e$  space where the group-speed is less than the phase-speed but still significant. Figure 5.8 also shows that the wavelength is shortest at the part of the structure furthest from the main wavepacket: since the shortest wavelengths (largest  $k_y\lambda_e$  components) have the slowest group-speeds they will naturally fall furthest behind, so this is to be expected.

We can also study the evolution of  $\tilde{N}$  at fixed  $\tilde{y}$ , as shown in Figure 5.9. An oscillatory character is readily apparent, with a period of approximately  $1.6\tau_{IM}$  ( $\omega \approx 0.6\omega_{IM}$ ) which is consistent with the IM-wave dispersion relation for  $k_y\lambda_e = 1$ . Although ionisation and recombination have been switched off, the disturbance at  $\tilde{y} = 0.6$  decays rapidly: this is because group-speeds remain significant, transporting energy in the direction of the electric field and causing the oscillations to

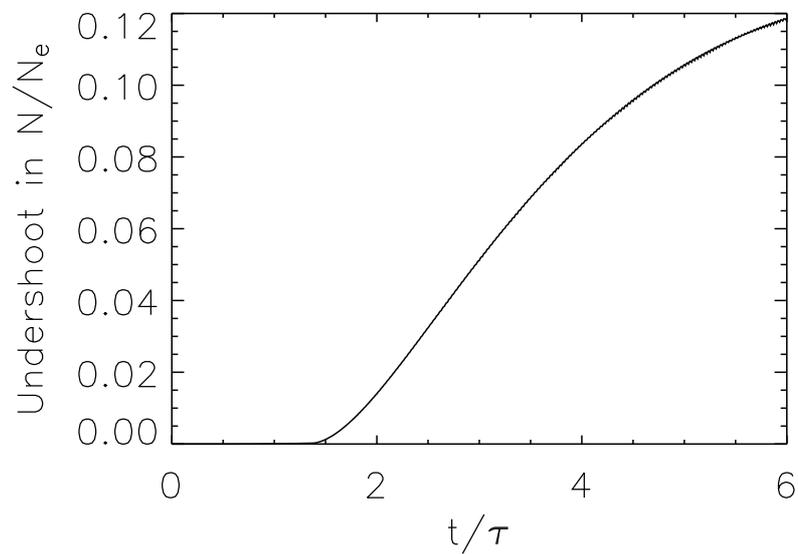


Figure 5.7: Growth of undershoot behind a non-linear wavepacket subject to electron inertial effects.

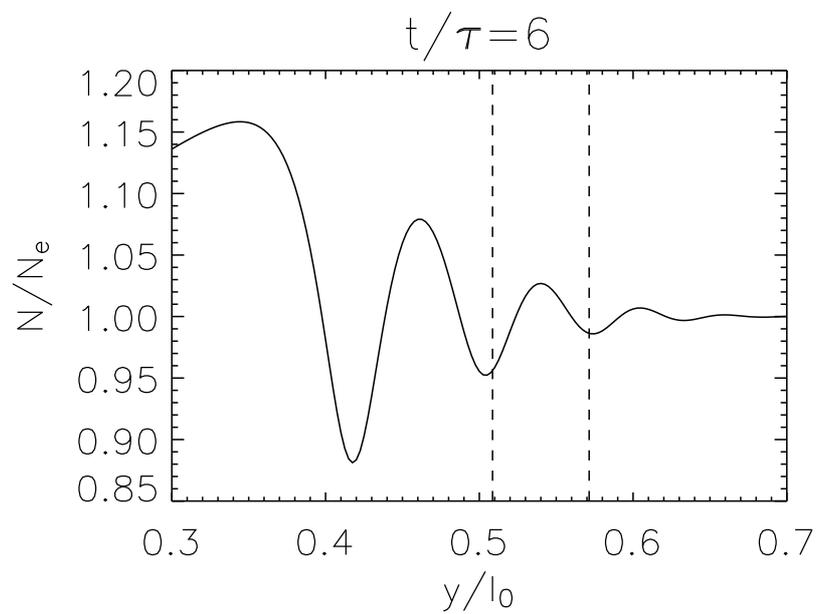


Figure 5.8: Close-up of small-scale waves behind the wavepacket at  $\tilde{t} = 6$ . These waves result from steepening of the trailing-edge of the wavepacket, and have a wavelength of approximately  $2\pi\lambda_e$  (indicated as the separation between the pair of vertical dashed lines).

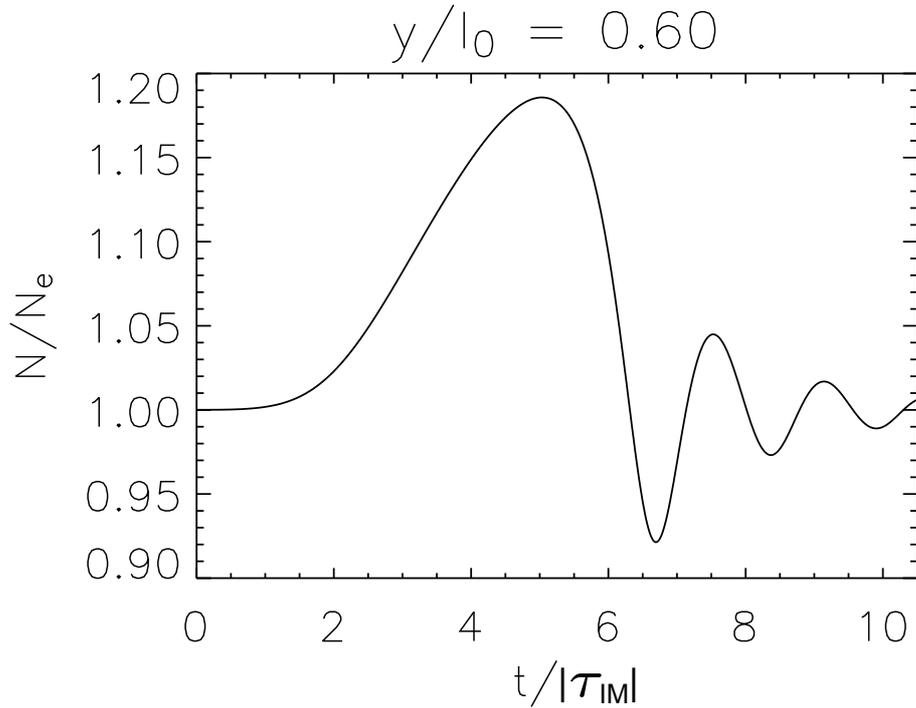


Figure 5.9: Evolution of  $\tilde{N}$  at  $\tilde{y} = 0.6$ . The plot shows the passage of the wavepacket over  $\tilde{y} = 0.6$ , followed by subsequent oscillations.

decay locally.

The origin of the small-scale, inertial waves can be understood in two ways. The first is to think about the effect of electron inertia on the discontinuity that forms in the ideal scenario. If a steep edge in  $N$  passes over a given location, the plasma-density at that location makes a rapid transition, dropping down towards the steady-state density. Where electron inertia is included in the magnetosphere, the transition is imperfect, and finite electron mass causes an undershoot, then an overshoot and so on. Therefore, the plasma-density oscillates, and continues to do so until either damped by ionisation and recombination, or energy is carried away in the direction of the electric field by the small group-speeds.

Alternatively, we can think of inertial IM-waves from the beginning, making use of the dispersion diagram (Figure 5.3). Initially almost all of the wavepower is at wavelengths much larger than the electron inertial length, so these advect in the ideal limit. This, however, causes steepening at the trailing-edge which transfers power to smaller wavelengths at the expense of the larger wavelengths. These smaller wavelengths have group-speeds less than the ideal advection speed, so as power is transferred to these wavelengths, the corresponding features fall behind the main wavepacket. Because they typically have a smaller amplitude than the main wavepacket, the trailing features are freed from the non-linear cascade to ever smaller wavelengths: significant power does not reach wavelengths much below the electron inertial length. This kind of behaviour is

known to occur in other non-linear systems with a source of dispersion, for example in viscous shocks (Hood, 2010), and in collisionless plasma shocks (Ofman et al., 2009).

## 5.5 Discussion

### 5.5.1 Typical Velocities, Frequencies and Decay Times

We have seen that ideal ionosphere-magnetosphere waves move in the direction of the background electric field at a speed,  $v_{IM}$ , given by (5.2). This characteristic speed is an important quantity, and one that can be written in terms of any one of the incident magnetic field perturbation ( $b_i$ ), incident electric field perturbation ( $E_i$ ), total magnetic perturbation ( $b_x^T$ ) or total electric field perturbation ( $E_y^T$ ). Using equations (3.38), (3.43), and (3.45), and introducing

$$M_P = \frac{\Sigma_P}{eN} \quad (5.53)$$

as the Pedersen mobility in the E-region,  $v_{IM}$  has the following forms:

$$v_{IM} = \frac{2\beta N_e b_i}{e\mu_0(N_e + \beta N)^2} = \frac{2M_P b_i}{\mu_0 \Sigma_A (1 + \Sigma_P/\Sigma_A)^2} \quad (5.54)$$

$$= \frac{2\Sigma_{P0} N_e E_i}{e(N_e + \beta N)^2} = \frac{2M_P E_i}{(1 + \Sigma_P/\Sigma_A)^2} \quad (5.55)$$

$$= \frac{N_e b_x^T}{e\mu_0 N(N_e + \beta N)} = \frac{M_P b_x^T}{\mu_0 \Sigma_A (1 + \Sigma_P/\Sigma_A)} \quad (5.56)$$

$$= \frac{\Sigma_{P0} E_y^T}{e(N_e + \beta N)} = \frac{M_P E_y^T}{(1 + \Sigma_P/\Sigma_A)}. \quad (5.57)$$

To estimate  $v_{IM}$  for ‘typical active conditions’, we might consider  $M_P = 10^4 \text{ m}^2\text{s}^{-1}\text{V}^{-1}$ ,  $E_y^T \approx 0.2 \text{ Vm}^{-1}$  and  $\Sigma_P/\Sigma_A \approx 10$ . Putting these into equation (5.57) gives a typical advection speed of  $180 \text{ ms}^{-1}$ . Note that this could be increased significantly by ionospheric depletion, which decreases the value of  $\Sigma_P/\Sigma_A$  on the denominator of (5.57).

The advection speed of IM-waves can be appreciated in context if we compare it to the ionospheric  $\mathbf{E} \times \mathbf{B}_0$  drift velocity,

$$v_E = \frac{|\mathbf{E} \times \mathbf{B}_0|}{B^2} \approx \frac{E_y^T}{B_0}. \quad (5.58)$$

Comparing these speeds,

$$\frac{v_{IM}}{v_E} = \frac{B_0 M_P}{(1 + \beta N/N_e)} = \frac{B_0 M_P}{(1 + \Sigma_P/\Sigma_A)}. \quad (5.59)$$

Here,  $M_P$  and  $B_0 = 5 \times 10^{-5}$  T are well determined quantities; the free quantity is  $\beta N/N_e = \Sigma_P/\Sigma_A$ , the ratio of Pedersen and ideal Alfvén conductances. Substituting for  $M_P$  and  $B_0$ ,

$$\frac{v_{IM}}{v_E} \approx \frac{0.5}{1 + \Sigma_P/\Sigma_A}.$$

For an undepleted ionosphere, equation (5.60) shows that  $v_{IM}$  is significantly smaller than  $v_E$ ; for example, putting  $\Sigma_P/\Sigma_A \approx 10$  into (5.60) gives  $v_{IM}/v_E \approx 0.05$ . In an E-region density cavity, however, it is possible to have  $\Sigma_P/\Sigma_A \ll 1$ , so that  $v_{IM}/v_E \approx 0.5$  (for the values of  $M_P$  and  $B_0$  assumed above), making  $v_{IM}$  a significant fraction of the  $\mathbf{E} \times \mathbf{B}_0$  velocity.

Electron inertial effects have been considered in this chapter, and the strongly inertial limit ( $k_y \lambda_e \gg 2\pi$ ) is characterised by oscillations at a characteristic angular frequency  $\omega_{IM}$  (5.33). This  $\omega_{IM}$  is based on a linear analysis, and can be expressed using any one of the incident magnetic field perturbation ( $b_i$ ), incident electric field perturbation ( $E_i$ ), equilibrium total magnetic field perturbation ( $b_{x0}$ ), or equilibrium total electric field perturbation ( $E_{y0}$ ). Useful arrangements are:

$$\omega_{IM} = \frac{2b_i}{e\mu_0 N_{SS} (1 + \beta N_{SS}/N_e) \lambda_e} = \frac{2M_P b_i}{\mu_0 \Sigma_{P,SS} (1 + \Sigma_{P,SS}/\Sigma_A) \lambda_e} \quad (5.60)$$

$$= \frac{2E_i}{e\mu_0 v_A N_{SS} (1 + \beta N_{SS}/N_e) \lambda_e} = \frac{2M_P E_i}{(\Sigma_{P,SS}/\Sigma_A) (1 + \Sigma_{P,SS}/\Sigma_A) \lambda_e} \quad (5.61)$$

$$= \frac{b_{x0}}{e\mu_0 N_{SS} (\beta N_{SS}/N_e) \lambda_e} = \frac{M_P b_{x0}}{\mu_0 \Sigma_{P,SS} (\Sigma_{P,SS}/\Sigma_A) \lambda_e} \quad (5.62)$$

$$= \frac{E_{y0}}{e\mu_0 v_A N_{SS} \lambda_e} = \frac{M_P E_{y0}}{(\Sigma_{P,SS}/\Sigma_A) \lambda_e}; \quad (5.63)$$

where  $N_{SS}$  is the steady-state value of  $N$ , which may be different from  $N_e$  in a local analysis if field-aligned current is present.

Arrangements of  $\omega_{IM}$  in terms of electric fields can be simplified further by expanding

$$e\mu_0 v_A \lambda_e = e\mu_0 \frac{B_0}{\sqrt{\mu_0 m^+ n_m}} \sqrt{\frac{m^-}{\mu_0 n_m e^2}} = \frac{B_0}{n_m} \sqrt{\frac{m^+}{m^-}}, \quad (5.64)$$

where  $n_m$  is the magnetospheric plasma-density. Putting this into equation (5.63) gives

$$\omega_{IM} = \frac{n_m}{B_0} \sqrt{\frac{m^+}{m^-}} \left( \frac{E_{y0}}{N_{SS}} \right). \quad (5.65)$$

Our model assumes a uniform magnetosphere, whereas physical values of  $n_m$  vary from about  $10^6 \text{ m}^{-3}$  in the equatorial plane of the magnetosphere to about  $10^{11} \text{ m}^{-3}$  in the F2 peak. Since IM-waves are an interaction between the magnetosphere and the E-region, it is likely that ‘magnetospheric’ conditions just above the E-region, at the bottom of the F-region, determine the value of  $\omega_{IM}$ . Therefore, a choice of oxygen for the ion mass, and a ‘magnetospheric’ number density

of  $n_m = 10^9 \text{ m}^{-3}$  seem reasonable first estimates to deduce typical values of  $\omega_{IM}$ . Substituting these into equation (5.65) with  $B_0 = 5 \times 10^{-5} \text{ T}$ ,  $E_{y0} = 0.2 \text{ Vm}^{-1}$  and  $N_{SS} = hn_{SS}$ , where  $n_{SS} \approx 10^{11} \text{ m}^{-3}$  for day conditions and  $h = 2 \times 10^4 \text{ m}$ , gives

$$\omega_{IM} \approx 0.69 \text{ rad s}^{-1},$$

which corresponds to a period

$$\tau_{IM} \approx 9 \text{ s}.$$

At night,  $n_{SS}$  can fall by a factor of ten due to reduced ionisation, and by another factor of ten in E-region plasma-density cavities; together, these decreases will reduce  $\tau_{IM}$  by a factor of one hundred. On the other hand, decreases in the electric field strength can increase  $\tau_{IM}$  several fold. Therefore,  $\tau_{IM}$  is likely to vary from several hundredths of a second to a few minutes. Note that this overlaps the reported 20–40 s periods of small-scale Alfvén waves observed in large-scale current systems (Karlsson et al., 2004).

Since IM-waves are damped by recombination and ionisation, it is also useful to estimate the decay time,  $\tau_d = 1/2\alpha n_{SS}$  (5.8). Here, we will take the recombination coefficient  $\alpha \approx 3 \times 10^{-13} \text{ m}^3\text{s}^{-1}$ . For a typical daytime number density  $n_{SS} \approx 10^{11} \text{ m}^{-3}$ , equation (5.8) gives  $\tau_d \approx 17 \text{ s}$ . This suggests that those IM-waves that observations resolve will be short-lived in the undepleted daytime E-region. Decreases in  $n_{SS}$  at night, and through E-region depletion, can increase  $\tau_d$  and may make it as large as  $\tau_d \approx 30 \text{ min}$  for a one hundred fold decrease in  $n_{SS}$  (not unreasonable). Therefore, under suitable conditions, IM-waves may last for many (tens or hundreds of) periods, or advect a significant distance in the case of an ideal wavepacket.

### 5.5.2 IM-Waves and Ionospheric Feedback Instability

To the best of my knowledge, this is the first study to have examined and characterised ionosphere-magnetosphere waves. Nonetheless, normal-mode dispersion relations for the coupled ionosphere-magnetosphere system have previously been published in the context of ionospheric feedback instability (IFI) (Atkinson, 1970; Sato and Holzer, 1973; Holzer and Sato, 1973; Sato, 1978). Linear IM-waves can be found in such results, albeit with less prominence than we have given them here. It is only proper, therefore, that we should comment on the close relationship between IM-waves and IFI, and highlight the presence of IM-wave solutions in previous studies.

### 5.5.2.1 IM-Waves Subject to Reflection of Upgoing Alfvén Waves Above E-region

In the presence of a large-scale background electric field, a small-scale perturbation in  $N$  produces small-scale field-aligned currents in the magnetosphere, which take the form of upgoing Alfvén waves. If upgoing Alfvén waves escape to infinity, then the small-scale FACs produce IM-wave behaviour, as showcased in this chapter. If, however, upgoing Alfvén waves are reflected back towards the E-region, from the conjugate ionosphere or the steep gradient in Alfvén speed that forms the ionospheric Alfvén resonator (IAR), then it also plays a secondary role that can lead to instability.

The secondary evolution of an IM-waveform (from trapping of Alfvén waves inside a ‘magnetospheric’ cavity) depends on the phase-difference with which small-scale Alfvén waves, arrive back at the E-region after being reflected from a point somewhere above it. One possibility is illustrated in Figure 5.10. Part (a) of Figure 5.10 illustrates the ‘primary’ FACs produced by an ideal perturbation in  $N$  with the form of a wavepacket; as we have seen previously, the primary FACs cause advection of the wavepacket. Part (b) illustrates the location of the wavepacket (which has moved in the direction of the electric field) at the moment when small-scale currents reflected at a point above the E-region arrive back at the E-region. For this illustration, we have chosen the Alfvén wave travel-time, from E-region to point of reflection back to E-region, so that the new upward current coincides with the peak of the wavepacket. The ‘secondary’ upward FAC deposits new electrons here, causing the wavepacket to grow; as the wavepacket grows, the primary FAC increases, which increases the secondary FAC at later times, and so on.

If the growth caused by secondary FACs exceeds decay by recombination, then the packet becomes unstable and grows to large amplitudes: this is the basis of IFI. The secondary currents also alter number density behind the original wavepacket, creating a series of peaks and troughs in  $N$ . Furthermore, as the disturbance increases in amplitude through IFI, it will enter the non-linear regime where  $v_{IM}$  noticeably varies with  $N$ ; this, in turn, will produce the wavebreaking behaviour that we studied in Section 5.4. Alternative phase-differences between  $N$  and the secondary currents can lead to quite different possibilities: these deserve detailed investigation in the future.

Whilst I have not, myself, performed simulations with partial reflection of upgoing Alfvén waves back to the E-region, Lysak and Song (2002) do present results for just such a scenario, which we reproduce in Figure 5.11 from their Figure 2c. In their paper, Lysak and Song (2002) identified the growth of waves with IFI, but the other features — propagation of the wavepacket, troughs catching up with crests, and the production of a wavetrain behind the original wavepacket — were simply described as ‘non-linear evolution of the feedback instability’. All of these additional features are exactly what we expect from an IM-waves interpretation. I therefore propose that it is more appropriate to describe these results as ‘evolution of IM-waves, subject to reflection

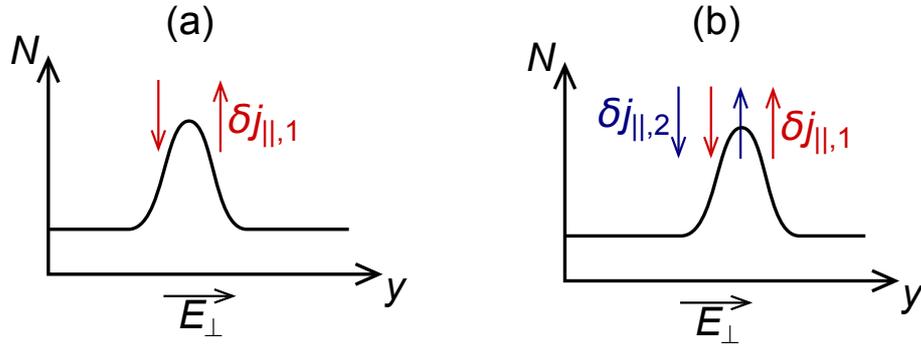


Figure 5.10: Action of ionospheric feedback instability on an ideal IM-wavepacket. (a) A perturbation in  $N$  produces small-scale FACs (red) in the magnetosphere which cause the wavepacket to advect in the direction of the electric field at speed  $v_{IM}$ . These ‘primary’ FACs propagate upwards as Alfvén waves. (b) It is possible that upgoing Alfvén waves are partially reflected somewhere above the E-region, returning to the E-region after some travel-time as ‘secondary’ FACs (blue). If  $v_{IM}$ , the Alfvénic travel-time and the spatial-scale of the wavepacket are such that an upward secondary-FAC aligns with the peak of the density perturbation, then the density perturbation will grow. This, in turn increases the strength of the small-scale FACs, and so on, leading to ionospheric feedback instability. Note, also, the secondary current arriving behind the wavepacket: this decreases  $N$ , beginning the formation of a wavetrain behind the original wavepacket.

of upgoing Alfvén waves’, since IFI forms only one part of this rich behaviour.

### 5.5.2.2 Ionospheric Feedback Instability as Resonance with IM-Waves

We have shown that coupling of the E-region and magnetosphere generates its own characteristic frequency of oscillation,  $\omega_r$  (5.32). This invites us to conjecture that ionospheric feedback instability can be interpreted as resonance between a magnetospheric cavity and IM-waves.

Various magnetospheric cavities have been proposed to be of relevance to IFI, but the fastest growth occurs for trapping in the ionospheric Alfvén resonator (Lysak, 1991) which typically has periods of a few seconds. In Section 5.3 we showed that  $\omega_r \leq \omega_{IM}$ , which equation (5.65) gives as

$$\omega_{IM} = \frac{n_m}{B_0} \sqrt{\frac{m^+}{m^-}} \left( \frac{E_{y0}}{N_{SS}} \right).$$

Estimates of  $\omega_{IM}$  (Section 5.5.1) correspond to periods from about a few hundredths of a second to a few minutes. Since  $\omega_{IM}$  is an upper limit on the frequency of IM-waves, IM-waves can only resonate with high-frequency IAR oscillations if  $\omega_{IM}$  is large: examining the form of  $\omega_{IM}$  above, this corresponds to large perpendicular electric field and low E-region number density (hence low Pedersen conductivity). The conditions favouring IFI, so obtained, are exactly those obtained from linear stability analysis (Sato, 1978; Lysak and Song, 2002).

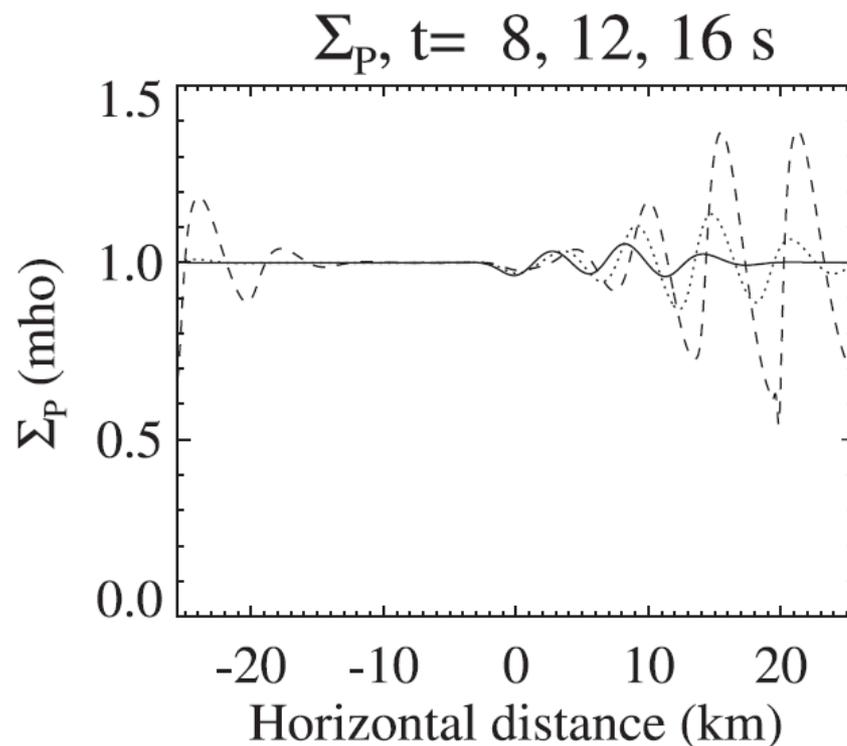


Figure 5.11: Reproduction of Figure 2c from Lysak and Song (2002), showing evolution of IM-waves, subject to reflection of upgoing Alfvén waves. Here  $\Sigma_P$ , which is proportional to  $N$ , is shown at three times,  $t = 8$  s (solid curve), 12 s (dotted curve) and 16 s (dashed curve). The simulation is initialised to generate a wavepacket in the centre of the domain, electric field is directed to the right, and boundary-conditions are periodic. The initial wavepacket propagates in the direction of the electric field, as expected from the theory of ideal IM-waves. Also note the formation of a wavetrain behind the original wavepacket, and steepening as the growing waveform enters the non-linear regime. Copyright 2002 American Geophysical Union. Reproduced by permission of American Geophysical Union.

In addition, resonance ideas lead us to suspect that the fastest growing mode of the instability is likely to be the normal-mode IM-wave whose frequency,  $\omega_r$ , matches the frequency of the magnetospheric cavity or a harmonic thereof. This expectation should be confirmed or disproved by future research into the link between IM-waves and IFI.

### 5.5.2.3 Linear I-M Dispersion Relation in Earlier Works

The linear (normal-mode) dispersion relation for ionosphere-magnetosphere coupling was first written down by Sato (1978), but the imaginary part (which corresponds to growth and damping) has received almost exclusive attention until now. In comparison, IM-wave solutions have received little attention, but they are nonetheless present in earlier works. As an example, let us consider the dispersion relation obtained by Lysak and Song (2002) (their equation (11)), which can be written as

$$\omega = \mathbf{k}_\perp \cdot \left( \mathbf{v}_E + \frac{\gamma \mathbf{v}_d}{1 + Z \Sigma_{P,SS}} \right) - 2i\alpha n_{SS}. \quad (5.66)$$

Here,  $\mathbf{v}_E$  is the  $\mathbf{E} \times \mathbf{B}_0$  drift velocity in the E-region,  $\gamma$  is a factor that gives the number of electrons produced in the E-region for every precipitating electron ( $\gamma = 1$  if FACs are carried by low-energy electrons),  $\mathbf{v}_d = M_P \mathbf{E}_\perp - M_H \mathbf{E}_\perp \times \hat{\mathbf{B}}_0$  is the relative drift velocity between ions and electrons in the E-region, and

$$Z = \mu_0 v_A \sqrt{1 + k_\perp^2 \lambda_e^2} \left( \frac{1 + R_c e^{2i\omega T}}{1 - R_c e^{2i\omega T}} \right) \quad (5.67)$$

is the magnetospheric impedance, assuming that upgoing Alfvén waves are reflected with reflection coefficient  $R_c$  after travelling upwards from the E-region for a time  $T$ . If upgoing Alfvén waves are free to escape to infinity (as in our model) then  $R_c = 0$  and equation (5.66) simplifies to

$$\omega = \mathbf{k}_\perp \cdot \left( \mathbf{v}_E + \frac{\gamma \mathbf{v}_d}{1 + \mu_0 v_A \Sigma_{P,SS} \sqrt{1 + k_\perp^2 \lambda_e^2}} \right) - 2i\alpha n_{SS}, \quad (5.68)$$

which is the normal-mode dispersion relation for IM-waves with a fully 2D sheet E-region.

Before discussing the properties of 2D IM-waves, let us show that the 1D limit of equation (5.68) recovers equation (5.30). Assuming invariance in the  $\mathbf{E} \times \mathbf{B}_0$  direction,  $\mathbf{k}_\perp \cdot (\mathbf{E} \times \mathbf{B}_0) = 0$  and equation (5.68) becomes

$$\omega = \frac{k_\perp \gamma M_P E_\perp}{1 + \mu_0 v_A \Sigma_{P,SS} \sqrt{1 + k_\perp^2 \lambda_e^2}} - 2i\alpha n_{SS} \quad (5.69)$$

$$\Rightarrow \omega = \gamma \left( \frac{M_P E_\perp}{\mu_0 v_A \lambda_e \Sigma_{P,SS}} \right) \left( \frac{\mu_0 v_A \Sigma_{P,SS} k_\perp \lambda_e}{1 + \mu_0 v_A \Sigma_{P,SS} \sqrt{1 + k_\perp^2 \lambda_e^2}} \right) - 2i\alpha n_{SS} \quad (5.70)$$

$$\Rightarrow \omega = \gamma \omega_{IM} \left( \frac{\mu_0 v_A \Sigma_{P,SS} k_\perp \lambda_e}{1 + \mu_0 v_A \Sigma_{P,SS} \sqrt{1 + k_\perp^2 \lambda_e^2}} \right) - 2i\alpha n_{SS}, \quad (5.71)$$

where

$$\omega_{IM} = \frac{M_P E_\perp}{\mu_0 v_A \lambda_e \Sigma_{P,SS}} \quad (5.72)$$

comes from equations (5.53) and (5.63). Noting that  $\mu_0 v_A \Sigma_{P,SS} \equiv \beta N_{SS}/N_e$  and setting  $\gamma = 1$  recovers equation (5.30) for 1D IM-waves, with all of the properties demonstrated in this chapter.

### 5.5.3 IM-Waves with Fully 2D Sheet E-Region

The normal-mode dispersion relation for IM-waves with a fully 2D sheet E-region (5.68) can be expressed simply as

$$\omega = \mathbf{k}_\perp \cdot \mathbf{v}_\perp - \frac{i}{\tau_d}, \quad (5.73)$$

where  $\tau_d = 1/2\alpha n_{SS}$  is the decay time due to recombination, introduced in equation (5.8), and  $\mathbf{v}_\perp$  has the form of a velocity. Introducing  $v_P$  as the component of  $\mathbf{v}_\perp$  in the direction of  $\mathbf{E}_\perp$ , and  $v_H$  as the orthogonal component in the direction of  $\mathbf{E}_\perp \times \mathbf{B}_0$ ,

$$v_P = \frac{\gamma M_P E_\perp}{1 + \sqrt{1 + k_\perp^2 \lambda_e^2 \Sigma_{P,SS}/\Sigma_A}}, \quad (5.74)$$

$$v_H = \frac{E_\perp}{B_0} - \frac{\gamma M_H E_\perp}{1 + \sqrt{1 + k_\perp^2 \lambda_e^2 \Sigma_{P,SS}/\Sigma_A}}. \quad (5.75)$$

In the ideal limit ( $k_\perp^2 \lambda_e^2 \ll 1$ ),  $v_P$  and  $v_H$  are independent of  $k_\perp$ , making ideal IM-waves non-dispersive. As in the 1D case, 2D IM-waves advect in the direction of the electric field at speed

$$v_{P,ideal} = \frac{\gamma M_P E_\perp}{1 + \Sigma_{P,SS}/\Sigma_A}, \quad (5.76)$$

which is equal to  $v_{IM}$  for  $\gamma = 1$ . The new behaviour that we pick up by going over to 2D, is advection in the  $\mathbf{E}_\perp \times \mathbf{B}_0$  direction at a speed

$$v_{H,ideal} = \frac{E_\perp}{B_0} - \frac{\gamma M_H E_\perp}{1 + \Sigma_{P,SS}/\Sigma_A}, \quad (5.77)$$

which does not produce a signature in 1D because of invariance in this direction. The first term on the right-hand side of (5.77) is the  $\mathbf{E}_\perp \times \mathbf{B}_0$  drift speed, and the second term is similar to  $v_{P,ideal}$ , but with  $M_H$  taking the place of  $M_P$  on the numerator. Therefore, in a frame of motion moving with the  $\mathbf{E}_\perp \times \mathbf{B}_0$  drift velocity relative to the background neutrals, ideal IM-waves advect in the direction of the electric field at speed  $\gamma M_P E_\perp / (1 + \Sigma_{P,SS} / \Sigma_A)$ , and in the direction of the neutral particles at a speed  $\gamma M_H E_\perp / (1 + \Sigma_{P,SS} / \Sigma_A)$ .

Unlike the treatment of ideal IM-waves in Section 5.2, this 2D analysis (based on normal-modes) is only valid for linear disturbances. It should be quite possible, however, to extend the non-linear approach to 2D, guided by the results of, for example, Lysak and Song (2002); Hasegawa et al. (2010) for handling of  $\mathbf{E}_\perp \times \mathbf{B}_0$  drift and Hall conductance. This task is left to the future, but will nonetheless be of significant interest. Further work is also needed to investigate the behaviour of 2D IM-waves in the strongly inertial regime.

#### 5.5.4 IM-Waves and Ionospheric Heaters

One method by which we can influence the ionosphere-magnetosphere system, is by using high-frequency (HF) radars to heat electrons in the E-region. This heating lowers the ion-electron recombination rate,  $\alpha$ , producing local enhancements in number density. In the context of IM-waves, these radars, called *ionospheric heaters*, open two exciting possibilities: (i) that we can use ionospheric heaters to experimentally verify IM-wave theory; and (ii) that IM-wave theory can help efforts to efficiently send Alfvén waves into the magnetosphere by modification of the E-region.

##### 5.5.4.1 Experimental Verification of IM-Wave Theory

First let us consider how a density perturbation, created by an ionospheric heater, would be expected to behave as an IM-wave. The ideal conditions for such an experiment are low background number density so that IM-waves have a long lifetime (see discussion of the decay time in Section 5.5.1), and a strong background perpendicular electric field so that IM-waves evolve significantly during their lifetime. These conditions naturally occur at night in the downward return current region adjacent to visible auroral arcs — the same conditions favoured by IFI experiments (Streltsov et al., 2010). If ionospheric heating creates an elliptical perturbation that is extended in the  $\mathbf{E} \times \mathbf{B}_0$  direction, then it will evolve as a 1D IM-wave, of the sort that we have explored extensively in this chapter.

Density perturbations, produced by heating, will evolve as ideal IM-waves provided they are much wider than the electron inertial length in the low-altitude magnetosphere. They should,

therefore, advect in the direction of the electric field at a speed given by  $v_{IM}$ , and may steepen as troughs catch up with crests, depending on the amplitude of the perturbations. Thus, an experimental programme might try the following types of driving, for which predictions are made:

1. Beam constantly on at a fixed location.

We expect to see E-region density increase at the heated location. Advection of ideal IM-waves will then create a tail, reminiscent of a wind-sock, and this will decay with an e-folding length (parallel to the electric field)  $\lambda_d = (v_{IM}\tau_d)^{-1}$ .

2. Single pulse at a fixed location.

A single pulse will create a wavepacket, which advects as an ideal IM-wave. Observations of the wavepacket after the heating pulse should reveal advection, damping, and (depending on the amplitude of the wavepacket) wavebreaking. Here, the aim is to experimentally reproduce the results of Sections 5.2.3 and 5.4.1.

3. Pulsed driver at a fixed location.

By pulsing the driver, or modulating it with a periodic function, a wavetrain can be produced, which again should exhibit the features of advection, decay by recombination, and possibly wavebreaking.

We can also imagine focusing the radar beam to an ellipse that is narrower than the electron inertial length in the magnetosphere. A wavepacket produced in this fashion should remain fairly stationary (in the direction parallel to the electric field), and oscillate at  $\omega_{IM}$ , reproducing the simulation presented in Section 5.3.3. Unfortunately, with present technology, it seems doubtful that an ionospheric heater could achieve the necessary tight focus (a few kilometres or less), or that observing instruments would reliably resolve these features. It nonetheless remains a possibility for the future.

#### 5.5.4.2 Using IM-Waves to Send Waves into the Magnetosphere

Whilst we have mostly viewed IM-waves in terms of E-region number density, they are equally associated with field-aligned currents. They can therefore be used to produce FACs which propagate upwards in the form of shear Alfvén waves. Depending on conditions, these waves may largely escape from the ionospheric Alfvén resonator, which opens possibilities to efficiently send ULF Alfvén waves into the magnetosphere.

The idea of heating the E-region to send ULF (or VLF) Alfvén waves into the magnetosphere is well established and has been confirmed by several experiments. Possible applications for such a technique are numerous, for example, ULF Alfvén waves have the potential to: interact

with energetic particles in the magnetosphere (Inan et al., 1985); move plasma along field-lines by the (non-linear) ponderomotive force (Streltsov and Lotko, 2008); ‘tag’ magnetic field-lines with a signature recognisable by spacecraft (Robinson et al., 2000), aiding the mapping of Earth’s magnetic field; or be used for ‘active’ magnetoseismology, where measurements of the waves’ propagation (e.g. travel time to the conjugate ionosphere) are exploited to reveal properties of the magnetosphere. One can even imagine that it might be possible to trigger geomagnetic substorms by sending a wave into a marginally-stable region of the magnetosphere-ionosphere system: whilst this last idea is very speculative, the concept of actively influencing the timing of storms is also very appealing, whether storms are triggered early to reduce their impact (e.g. when less energy has accumulated) or to make the best use of scientific resources (e.g. to coincide with satellite orbits).

How do we get the biggest effect (strongest FAC) for a given heating power? The properties of IM-waves suggests an answer. Recall that an ionospheric heater creates a localised region where the recombination coefficient has been lowered. In the absence of a background electric field, the low recombination coefficient produces an E-region density enhancement that remains stationary and decays when heating is switched off; this does not create FACs because there is no incident magnetic field perturbation to reflect. In the presence of a background electric field, however, FACs are created, and the initial density enhancement advects as an IM-wave. Therefore, if the beam remains stationary, then the initial density perturbation will advect into unheated areas where the recombination coefficient is larger, damping the IM-wave. The damping problem can be avoided, however, if the heating region tracks at the IM-advection speed  $v_{IM}$ . This way, the density enhancement remains inside the region of suppressed recombination at all times, ensuring that it has the greatest possible amplitude, with correspondingly strong FACs.

We can also use ionospheric feedback instability to further increase the E-region density perturbation, and hence boost the strength of the upgoing FACs, by carefully choosing the width of our heating region. Figure 5.10 illustrates the desired situation for a Gaussian heating profile: IFI will cause growth if the wavepacket advects a distance  $\text{FWHM}/2$  in the time,  $T$ , taken for an Alfvén wave to complete on full cycle inside the magnetospheric cavity. We therefore wish to choose the width of the heating region so that

$$\text{FWHM} = \frac{2T}{v_{IM}}. \quad (5.78)$$

This analysis is in excellent agreement with recent results presented by Streltsov and Pedersen (2010), who used computer simulations to search for effective methods to produce FACs using ionospheric heating. These authors observed that E-region density features naturally move in the direction of the electric field (at a speed of  $74.3 \text{ ms}^{-1}$  in their simulation), and that stronger FACs are produced when the region of ionospheric heating is tracked at this speed. By linking their

---

results with an IM-wave interpretation, we are able to propose a formula for the advection speed ( $v_{IM}$ ), and to explain why efficiency improves if the heating region tracks at this speed, adding significant value to their findings. Recent correspondence has confirmed that  $74.3 \text{ ms}^{-1}$  agrees with the value of  $v_{IM}$  computed for their simulations, with  $v_A$  evaluated at the base of the F-region (they include a fully stratified magnetosphere).

# Coupled Magnetosphere-Ionosphere Dynamics Driven by Field-Aligned Currents

---

## 6.1 Introduction

This chapter seeks to answer, “How does the ionosphere react to a large-scale current system, given that the ionosphere also modifies magnetospheric current?” The ground has been carefully prepared: self-consistent models and codes have been developed (Chapter 3); steady-states have been closely examined (Chapter 4); and ionosphere-magnetosphere waves have been studied and developed as a means of understanding ionosphere-magnetosphere dynamics (Chapter 5). Now we apply these tool to ionosphere-magnetosphere dynamics when the system is driven by large-scale field-aligned currents originating from the outer magnetosphere.

The chapter is organised in three main sections, with the first division according to sustainability of the downward current density: the steady-states of Chapter 4 show a changeover when the current density drawn by the magnetosphere exceeds that which the ionosphere can supply by ionisation, so we treat the sustainable case,  $|2j_i| < j_c$  for all downward current, in its own Section 6.2. When  $|2j_i| > j_c$  in a downward current region, the dynamics become significantly more complex, with the formation and broadening of an ionospheric plasma-density cavity. It is useful to first examine the dynamics of broadening for an ideal magnetosphere, which we treat in Section 6.3, establishing key ideas before bringing in inertial effects in Section 6.4. As usual, the different sections are tied together in a discussion to conclude the chapter.

## 6.2 Dynamics for ‘Sustainable’ Current Densities

### 6.2.1 Numerical Case Study

To demonstrate how the E-region copes with a sustainable current density ( $|2j_i| < j_c$  in downward current regions) we can look at a numerical case study. Inertial length-scales are not expected to develop, so the simulation is performed with the ideal code described in Section 3.4, reusing the

setup described in Chapter 4 for analysis of steady-states (see Section 4.2.1 for details). For the physical parameters used in the simulation, the maximum current density that ionisation in the E-region can support is

$$j_c = \frac{\alpha N_e^2 e}{h} \approx 3.46 \times 10^{-6} \text{ A.} \quad (6.1)$$

Hence, an initial total current (incident plus reflected) of  $j_{\parallel 0} = 2 \times 10^{-6} \text{ A}$  is sustainable, and we use this to set  $\tilde{\alpha}$  in the simulation.

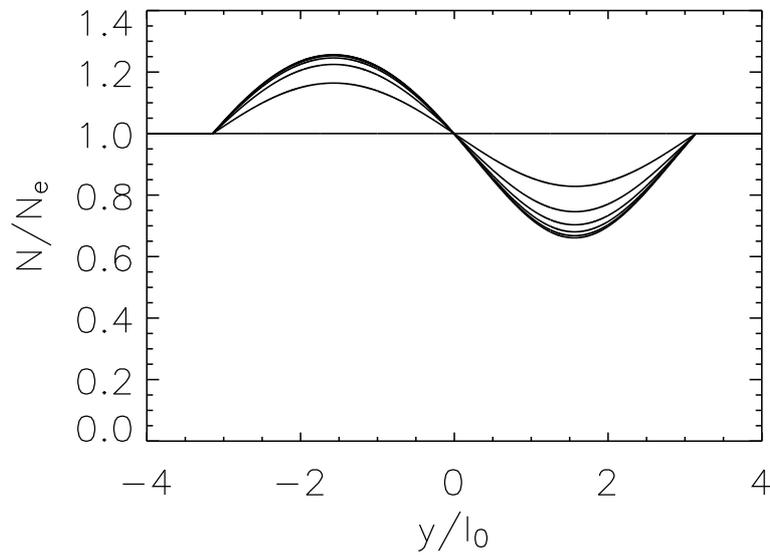


Figure 6.1: Evolution of  $\tilde{N}$  for a sustainable current density ( $j_{\parallel 0} = 2 \mu\text{Am}^{-2}$ ) and  $\beta = 100$ . Snapshots are plotted at intervals of  $\Delta\tilde{t} = 25$  between  $\tilde{t} = 0$  and 150.

Results are shown in Figure 6.1. At  $\tilde{t} = 0$ , the plasma-density is  $\tilde{N} = 1$  everywhere. Field-aligned current immediately alters the balance between loss and production processes: upward field-aligned current deposits electrons in the E-region, increasing  $\tilde{N}$  between  $\tilde{y} = -\pi$  and  $\tilde{y} = 0$ ; downward field-aligned current removes electrons from the E-region, so  $\tilde{N}$  falls between  $\tilde{y} = 0$  and  $\tilde{y} = \pi$ . Initially, the adjustment is rapid, but slows as the balance re-establishes itself; hence, at later times,  $\tilde{N}$  gradually relaxes to the upper steady-state solution.

### 6.2.2 Analytic Solution

Sustainable current densities ( $|2j_i| > j_c$  in downward current regions) are an important case, because they allow us to find a time-dependent solution to governing equation (3.46). Density cavities (regions where the ionospheric reflection coefficient varies significantly from  $r = -1$ ) do

not form if the E-region can easily supply the downward magnetospheric current by ionisation. Thus, throughout the evolution of the ionosphere,  $\beta N \gg N_e$ . In this limit, equation (3.46) simplifies to

$$\frac{\partial N}{\partial t} + \frac{2}{\mu_0 e} \frac{db_i}{dy} = \frac{\alpha}{h} (N_e^2 - N^2) \quad (6.2)$$

$$\Rightarrow \frac{\partial N}{\partial t} = \left( \frac{\alpha}{h} N_e^2 - \frac{2}{\mu_0 e} \frac{db_i}{dy} \right) - \frac{\alpha}{h} N^2. \quad (6.3)$$

Writing  $N = N_{upper}$  for the upper steady-state solution defined in (4.11),

$$\frac{\partial N}{\partial t} = \frac{\alpha}{h} (N_{upper}^2 - N^2) \quad (6.4)$$

which can be solved by separation of variables and partial fractions. The initial condition,  $N = N_e$  at  $t = 0$ , determines the constant of integration, yielding the solution

$$N = N_{upper} \left( \frac{(N_e + N_{upper}) \exp(2\alpha N_{upper} t/h - N_{upper} + N_e)}{(N_e + N_{upper}) \exp(2\alpha N_{upper} t/h) + N_{upper} - N_e} \right). \quad (6.5)$$

A plot of snapshots of the analytic solution is indistinguishable from Figure 6.1. The approximation  $\beta N \gg N_e$  is weakest at the location of minimum plasma-density, but even here, where disagreement between the two solutions is greatest, the largest difference in  $\tilde{N}$  is only 0.00723.

## 6.3 Dynamics for ‘Unsustainable’ Current Densities with Ideal Magnetosphere

### 6.3.1 Numerical Case Study

If the field-aligned current initially drawn by the magnetosphere cannot be met by ionisation in the E-region, then a density cavity forms and broadens (Section 4.2.2). Figure 6.2 shows the dynamics of broadening, using an ideal numerical simulation with  $\beta = 100$  and  $j_{||0} = 5 \mu\text{Am}^{-2}$  for data.

Evolution of the upward current channel is qualitatively identical to the case  $|2j_i| < j_e$ , following the course prescribed by equation (6.5). This is because all upward current, no matter how strong, is ‘sustainable’ because the magnetosphere acts as an essentially infinite source of electrons. The only exception is where the downward current channel eats in as it broadens.

At early times ( $\tilde{t} \leq 200$  in Figure 6.2), E-region plasma-density in the downward current channel evolves along familiar lines: the downward current increases loss of electrons, so the plasma-density decreases, and decreases rapidly for strong current densities. Around  $\tilde{t} = 200$ , an

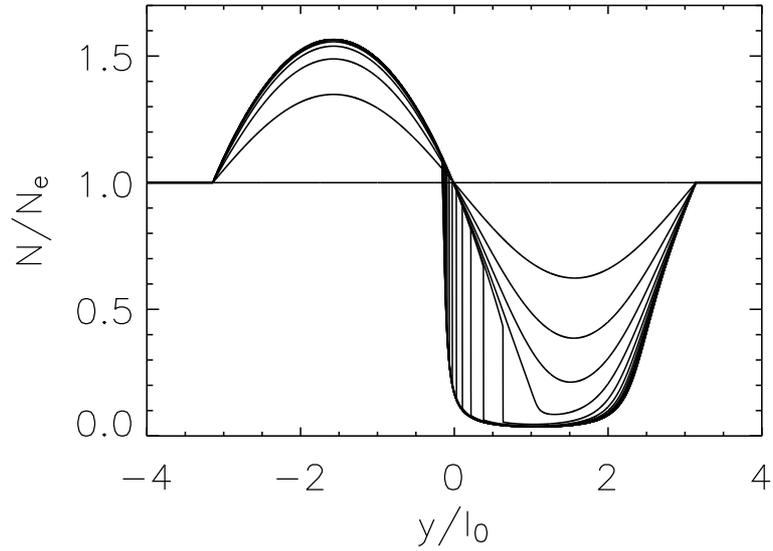


Figure 6.2: Evolution of  $\tilde{N}$  for an unsustainable current density ( $j_{\parallel 0} = 5 \mu\text{Am}^{-2}$ ),  $\beta = 100$  and an ideal magnetosphere. Snapshots are plotted intervals of  $\Delta\tilde{t} = 50$  between  $\tilde{t} = 0$  and 850.

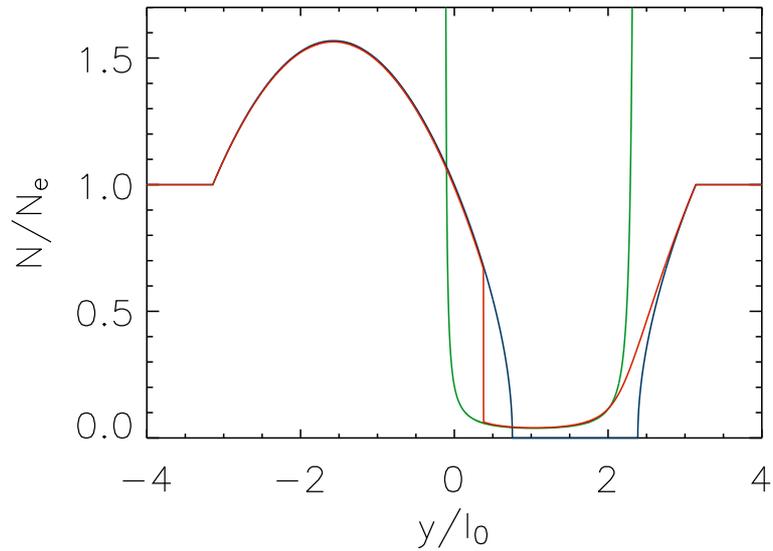


Figure 6.3: Relationship between the discontinuity and the upper and lower steady-state solutions. Here the red line shows  $\tilde{N}$  at  $\tilde{t} = 300$ , the blue line shows the upper steady-state, and the green line shows the lower steady-state. The discontinuity is a step transition from the upper solution to the lower solution.

asymmetry can be seen on the left-hand side of the downward current channel, and this rapidly develops into a propagating discontinuity that moves in the direction of the background electric field.

At the discontinuity,  $\tilde{N}$  appears to jump from one steady state to another Figure 6.2. Examining this further, Figure 6.3 plots the numerical solution at  $\tilde{t} = 300$ , together with the upper and lower steady-state solutions obtained in Chapter 4. It clearly shows that the discontinuity connects the upper and lower steady-states.

### 6.3.2 Analysis In Terms of Ideal IM-Waves

In the first stage of dynamic evolution, field-aligned current alters the balance between loss and production of electrons in the E-region, and the plasma-density adjusts accordingly as it seeks a new steady-state. One must, of course, ask which steady-state will the curve move towards? In most of the domain, the curve first encounters the upper steady-state and is quite happy to stay there. Exceptions are the boundary-layer at the edge of the downward current channel farthest from the upward channel, where the curve settles onto the matched steady-state obtained in Section 4.4, and the bottom of the downward current channel, where it meets a valid lower steady-state. There is one region, between the upward and downward current channels, where the plasma-density pauses on the upper steady-state, but there is also a perfectly valid lower steady-state underneath.

The second part of the evolution is the formation and travel of a propagating discontinuity. This feature is clearly related to the wavebreaking of ideal IM-waves (examined in Section 5.4.1), a process that inevitably occurs for non-linear IM-waves because plasma-density disturbances advect at a speed which depends on height-integrated number density,  $N$ , (5.2). Consequently, regions of low  $N$  ‘catch up’ with regions of high  $N$ , forming a discontinuity.

Where is advection in a steady-state if the curve appears stationary? Here, the image of a meandering river can be helpful. Picture a river from above: the river appears stationary, even though water flows along it. This is possible, because as water flows from the hills to the ocean, it experiences a changing push to left or right, which keeps it following the river. The steady-state is similar: it appears stationary, even though ‘density-curve’ constantly flows along it. Like water flowing towards the sea, the density-curve flows in the direction of the background electric field; and like water following a river, the density-curve follows the steady-state up and down as it experiences a changing balance of field-aligned current, ionisation and recombination (5.3).

Mathematically, this can be seen from the steady-state version of equation (5.1):

$$v_{IM} \frac{\partial N_{SS}}{\partial y} = F \equiv \frac{dN}{dt}. \quad (6.6)$$

Here,  $v_{IM}$  and  $F$  control the ‘flow’ of the density curve as points  $(y, N)$  move along characteristic trajectories. A point  $(y, N)$  moves a distance of  $\delta y = v_{IM}\delta t$  in the  $y$ -direction during elemental time-interval  $\delta t$ ; as it moves in  $y$ , it also moves up and down according to  $dN/dt \equiv F$ , changing by  $\delta N = F\delta t$ . It follows that the gradient of the characteristic trajectory in  $(y, N)$  space is

$$\frac{\delta N}{\delta y} = \frac{F\delta t}{v_{IM}\delta t} = \frac{F}{v_{IM}} = \frac{\partial N_{SS}}{\partial y}. \quad (6.7)$$

Therefore, as points on the curve moves along their characteristic trajectories, they follow the steady-state curve. This reinforces the idea of constant motion in the steady-state, even though  $N$  is unchanging in time at any fixed location. Figure 6.4 illustrates this interpretation.

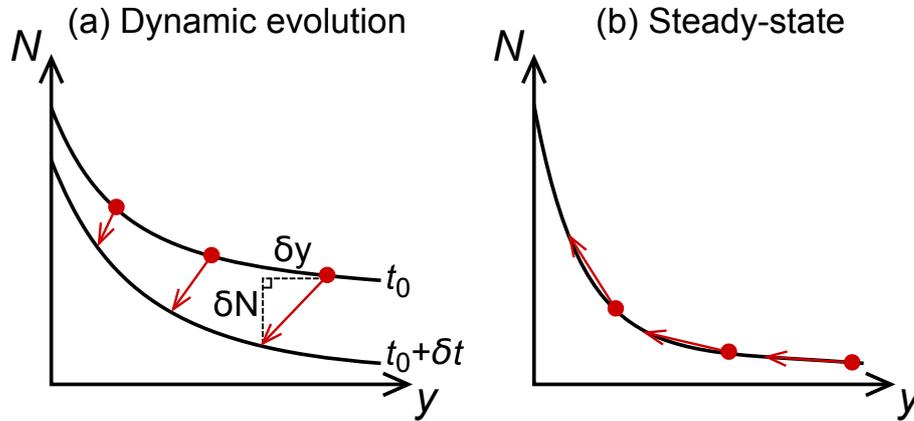


Figure 6.4: Motion of points  $(y, N)$  along characteristic trajectories. (a) As the density-curve evolves, the points  $(y, N)$  that make up this curve move along characteristic trajectories. The motion is controlled by  $v_{IM} = \delta y/\delta t$ , which governs motion in  $y$ , and  $F = \delta N/\delta t$ , which governs motion in  $N$ . (b) If characteristic trajectories follow the density-curve, then  $N$  is unchanging in time at any fixed location. This corresponds to the steady-state solution.

If the first stage of dynamic evolution populates parts of the upper and lower steady-states as described above, and points making up the density-curve move along steady-states, what happens next? The most interesting region lies within the downward current channel. At the centre of the current channel, the lower steady-state is populated, and this part of density-curve flows in the direction of the electric field. By so doing, it comes to a critical point where it undercuts a valid upper solution, but stays on its own course. Meanwhile, density-curve is flowing along the upper solution, but because the advection speed depends on  $N$ , the upper solution flows more slowly than the lower solution, and the lower solution catches up. The wave breaks, a discontinuity forms, and this travels in the direction of the electric field at the geometric mean of advection speeds on either side of the discontinuity (5.52). Eventually, the discontinuity approaches the singularity in the lower solution, and the difference in  $N$  across the discontinuity becomes ever smaller until it vanishes to leave the global steady-state everywhere.

## 6.4 Dynamics for ‘Unsustainable’ Current Densities with Electron Inertial Effects

Although the ideal analysis provides key insights into this system, it cannot have the final word. An ideal analysis assumes large length-scales, yet it results in the collapse of the perpendicular length-scale to zero, contradicting that initial assumption: once more, we need to include electron inertia in the magnetosphere.

### 6.4.1 Overview of Numerical Simulation

To investigate the influence of electron inertia in ionosphere-magnetosphere dynamics driven by large-scale field-aligned currents, we use the computer code described in Section 3.5. For direct comparison with the ideal simulation presented in Section 6.3.1, the system is driven with a similar incident wave: once again,  $j_i$  is sinusoidal, producing a single channel of upward current and an adjacent channel of downward current; and the ratio of greatest  $j_i$  to  $j_e$  is preserved across the two (ideal and inertial) simulations. Some differences, however, cannot be avoided. The most significant of these is that we are forced to perform this run at a lower value of  $\beta$  than in the ideal case, using  $\beta = 10$  to preserve numerical stability at the ionospheric boundary. I shall explain the reason for this — and discuss some consequences of lower  $\beta$  — after a first look at the simulation results. Also, remember that ideal and inertial simulations have different normalisations, so normalised times are not the same here as in Figure 6.2. Finally, this simulation includes the effects of electron inertia in the magnetosphere, using an electron inertial length  $\tilde{\lambda}_e = 0.025$ .

Simulation results are shown in Figure 6.5, which shows snapshots of  $\tilde{N}$  at six evenly spaced times, as well as  $\tilde{j}_z$  in the (very) low-altitude magnetosphere. At  $\tilde{t} = 3$ , the incident Alfvén wave is just about to reach the top of the E-region (at  $\tilde{z} = 0$ ), so  $\tilde{N}$  is undisturbed. Shortly afterwards, the incident Alfvén wave reflects from the E-region, reflection almost doubling the field-aligned current drawn by the magnetosphere. Once this current has connected with the E-region, it alters the balance between loss and gain of electrons in the E-region, and  $\tilde{N}$  changes in response. The snapshot at  $\tilde{t} = 7.5$  shows early evolution and strongly resembles the situation for ideal magnetosphere: this is reasonable since transverse length-scales ( $y$ -direction) are much greater than the electron inertial length at this stage.

In the ideal case, we saw an asymmetry develop in the downward current channel, as the side adjacent to the upward current channel steepened to form a travelling discontinuity. The inertial equivalent of the ideal discontinuity appears in the snapshots taken at  $\tilde{t} = 12$  and 16.5. If  $\beta = 10$ , the gap between the upper and lower solutions is considerably less than if  $\beta = 100$ , so even in ideal

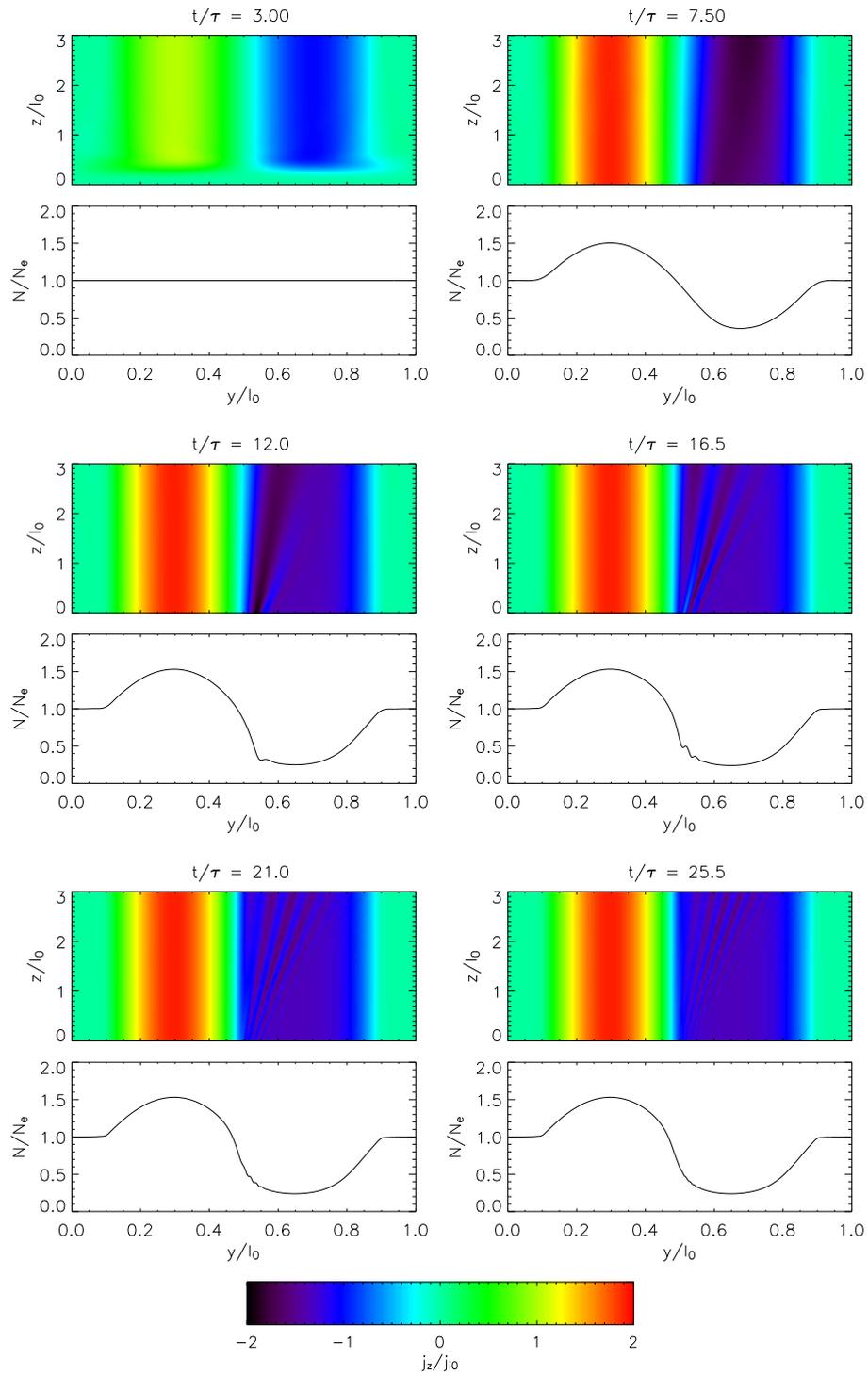


Figure 6.5: Evolution of  $\tilde{N}$  for an unsustainable current,  $\beta = 10$  and including electron inertial effects in the magnetosphere. The lower part of each panel plots height-integrated number density,  $\tilde{N}$ , while colour in the top part of each panel represents field-aligned current ( $\tilde{j}_z$ ) in the magnetosphere ( $y, z$ ). The ionosphere is located at  $\tilde{z} = 0$ .

simulations the discontinuity is less pronounced; also, electron inertia smooths out discontinuities in  $\tilde{N}$ , so the inertially-smoothed discontinuity can be a little tricky to spot at first. However, if one examines the panel of Figure 6.5 taken at  $\tilde{t} = 12$ , the discontinuity-equivalent is there to be seen in the form of a steep gradient between  $\tilde{y} = 0.5$  and  $\tilde{y} = 0.55$ . A clearer tell-tale sign is the intense concentration of downward current at the base of the magnetosphere, visible at the same  $\tilde{y}$ . (In the ideal case with  $b_i \neq 0$ , a discontinuity in  $\tilde{N}$  corresponds to a current sheet.)

The situation is reminiscent of the effects of electron inertia on wavebreaking of a non-linear wavepacket (Section 5.4.2). Sure enough, as well as smoothing out the discontinuity, electron inertia leads first to an undershoot in  $\tilde{N}$  ( $\tilde{t} = 12$  in Figure 6.5) and then a whole series of undershoots and overshoots ( $\tilde{t} = 16.5$  in Figure 6.5). Thus, broadening of the downward current channel acts as a source of inertial ionosphere-magnetosphere waves, which fill the downward current channel in the magnetosphere with upward propagating inertial Alfvén waves. These upward propagating inertial Alfvén waves take their energy from the large-scale incident Alfvén wave, energy being conserved by corresponding decreases in ionospheric heating; therefore, the magnetospheric waves do not cause the IM-waves on the ionosphere-magnetosphere boundary to decay.

By late times ( $\tilde{t} = 21$  and 25.5) broadening is complete, and there is no new source for small-scale IM-waves (in the form of a smoothed-discontinuity). Those small-scale IM-waves that already exist at the ionosphere-magnetosphere boundary continue to oscillate, and they develop an ever decreasing wavelength, but do not noticeably move from the region over which broadening occurred. Unreplenished by new waves, they decay by ionisation and recombination, as  $\tilde{N}$  relaxes to the steady-state.

#### 6.4.2 Limitation of $\beta$ in Simulation

Having described the main features seen in simulation results, I can now explain why we are forced to use a comparatively low value of  $\beta$  to run the simulation. In Section 3.5.7, it was noted that the ionospheric boundary-condition used by this code is stable for  $\beta\tilde{N} \gtrsim 0.2$ . At first, it seems that large  $\beta$  must be favourable for stability, but then one remembers that the minimum value of  $\tilde{N}$  in the steady-state is proportional to  $1/\beta$  (see Sections 4.2.3 and 4.3.4), so if we were dealing with ideal evolution, stability would be independent of  $\beta$ .

Inertial simulations, however, are subject to an extra consideration: the ‘broadening-front’ is followed by an undershoot in  $\tilde{N}$ , and this undershoot can be deep enough to push the simulation into instability. Experience has shown that evolution is qualitatively identical for different values of  $\beta$  but the undershoot is deeper for larger values of  $\beta$ , making the simulation unstable for  $\beta \gtrsim 10$ . This is why we are forced to use  $\beta = 10$  to follow the dynamics to the very end.

### 6.4.3 Inertial IM-Waves in Downward Current Channel

The production of inertial ionosphere-magnetosphere waves, and their subsequent evolution, merits further attention. The small-scale waves are most easily studied if we subtract the final steady-state solution from data  $\tilde{N}(\tilde{y}, \tilde{t})$ , the difference being the wave perturbations.

Figure 6.6 shows a close-up of the perturbation in height-integrated number density,  $\tilde{N}_{pert}$ , at two different times after the waves have been established:  $\tilde{t} = 18$  (red) and  $\tilde{t} = 24$  (blue). The IM-wave disturbance is fairly stationary (in terms of group velocity), damps in time, and has a sinusoidal structure in  $\tilde{y}$  with a wavelength that gets shorter as the simulation progresses. The electron inertial length is indicated in Figure 6.6 as the distance between the vertical dashed lines, and we can see that the diminishing wavelength passes through the electron inertial length just before  $\tilde{t} = 18$ , a short time after the IM-waves are first established.

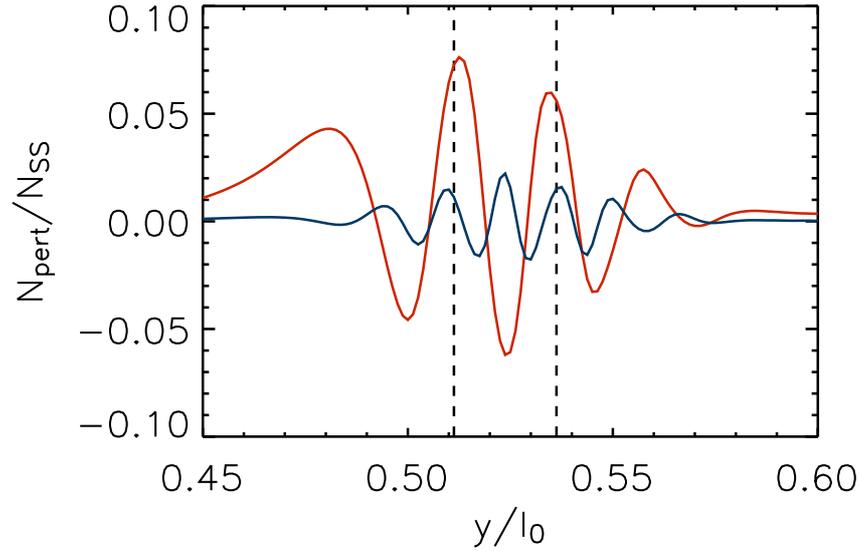


Figure 6.6: Perturbation in height-integrated number density at  $\tilde{t} = 18$  (red) and  $\tilde{t} = 24$  (blue).

The IM-waves are not just inertial: for most of their existence they are *strongly* inertial, with  $k_y \lambda_e \geq 2\pi$ . Referring to results of Section 5.3 (the dispersion diagram shown as Figure 5.3 is a particularly useful reference) we therefore expect these waves to oscillate with angular frequency  $\omega_{IM}$  given by equation (5.33). This expectation is confirmed by Figure 6.7, which shows  $\tilde{N}_{pert}$  at  $\tilde{y} = 0.54$  plotted against a time axis that has been renormalised by the local value of  $|\tilde{\tau}_{IM}|$ . As the simulation progresses, the period of oscillation does indeed tend to  $|\tilde{\tau}_{IM}|$ . It is also important to recognise that we are using the local value of  $\omega_{IM}$ , which varies significantly with  $y$ . Figure 6.8 plots  $\tilde{\omega}_{IM}$  as a function of position: it contains strong gradients, and varies by about a factor

seven across the region populated by small-scale IM-waves.

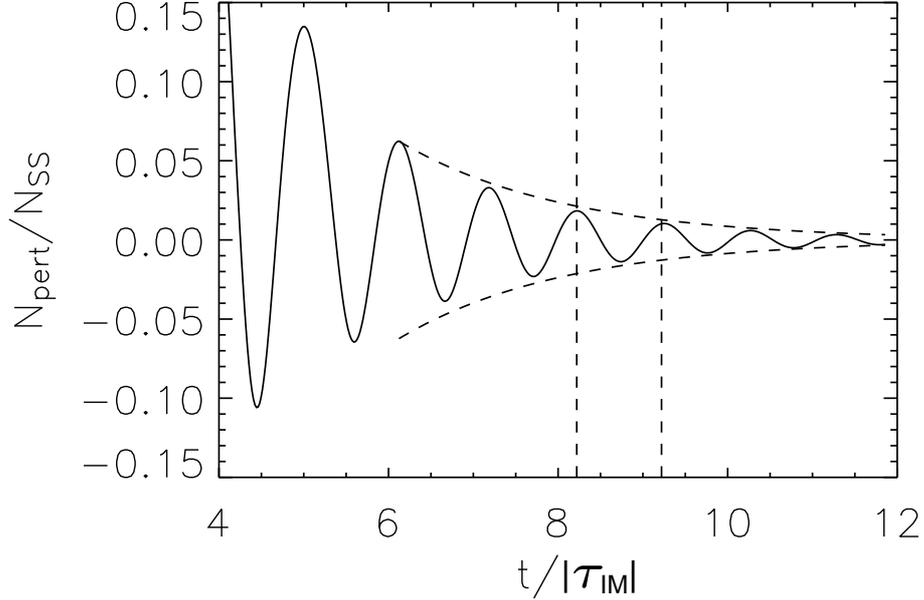


Figure 6.7: Perturbation in height-integrated number density at  $\tilde{y} = 0.54$ . The period of oscillation tends to  $|\tilde{\tau}_{IM}|$  at later times (e.g. the vertical dashed lines are separated by  $\tilde{t}/|\tilde{\tau}_{IM}| = 1$ ). The dashed envelope indicates exponential decay at later times, due to ionisation and recombination.

Now let us try to construct the full story of the small-scale, inertial IM-waves. First, the waves are produced by broadening of the downward current channel. Early evolution of the E-region should be understood in terms of ideal IM-waves, which cause one edge of the downward current channel to steepen (see Section 6.3.2). This steepening is similar to non-linear steepening of a wavepacket, which we considered in Section 5.4.2, where an example showed that in place of ideal discontinuity, inertial effects create a series of short-wavelength undershoots and overshoots behind a travelling ‘rapid transition’ in  $N$ .

The present scenario and the wavepacket example of Section 5.4.2 differ in the wavelength of the smallest scales (down to  $\tilde{\lambda}_e/2$  when driven by field-aligned current, compared with about  $2\pi\tilde{\lambda}_e$  in the wavepacket example). This, in turn, gives rise to a shorter period (because  $T \rightarrow |\tau_{IM}|$  as  $k_y\lambda_e \rightarrow \infty$ , where  $\tau_{IM}$  is given by equation (5.37)) and correspondingly small group-speeds (because  $v_g \rightarrow 0$  as  $k_y\lambda_e \rightarrow \infty$ ). Why do wavelengths differ so much between the two simulations? The answer lies in the steep gradients of  $\tilde{\omega}_{IM}$ , produced by the large-scale system of field-aligned current as a result of the non-uniform steady-state: because of these frequency gradients, the inertial IM-waves produced by broadening are subject to phase-mixing, of a similar nature to the phase-mixing of Alfvén waves seen in Chapter 2. Therefore, once created, their wavelength rapidly decreases (as seen in Figure 6.6), moving further and further into the strongly inertial regime  $k_y\lambda_e \gtrsim 2\pi$ , with its attendant properties.

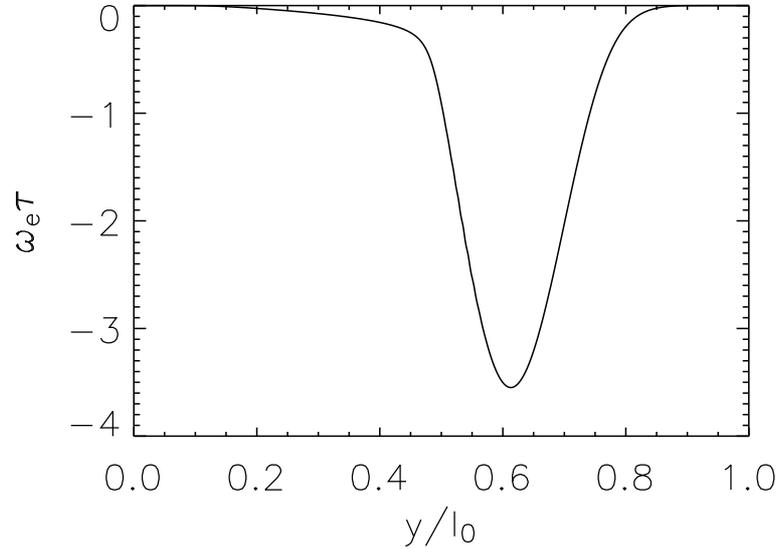


Figure 6.8: Plot showing  $\tilde{\omega}_{IM}$  as function of position. Broadening produces inertial IM-waves over  $\tilde{y} = 0.49$  to  $0.57$ , where there is a steep gradient in  $\tilde{\omega}_{IM}$ . This leads to phase-mixing, which shortens the wavelength over time.

Ultimately, the inertial IM-waves damp under the action of ionisation and recombination. At early times, the wavelength is sufficiently large that the group-speed is non-zero and energy can move in the direction of the electric field. At late times, however,  $k_y \lambda_e \gtrsim 2\pi$ , so the group speed goes to zero and changes in amplitude are due to damping, not movement of energy. We therefore expect exponential decay with a damping time,  $\tau_d$ , given by equation (5.31). A corresponding envelope is plotted on Figure 6.7 and the agreement is excellent.

## 6.5 Discussion

This study of ionosphere-magnetosphere dynamics has addressed two problems: (i) “How can we understand broadening of downward current channels and E-region density cavities?” and (ii) “What is the origin of intense small-scale Alfvén waves in large-scale systems of field-aligned current?” In both cases, the theory of IM-waves (outlined in Chapter 5) provides an excellent interpretive framework.

In Section 6.3, it was shown that downward currents cannot be sustained by the E-region if  $|2j_i| > j_c$  in the downward current region, because strong current densities remove electrons from the E-region more rapidly than they can be produced by ionisation. The downward current therefore has to broaden to reduce the current density. Since the system evolves as IM-waves,

broadening must occur in the direction of the electric field, and the widening rate will correspond to a discontinuity speed  $U = \sqrt{v_{IM}^+ v_{IM}^-}$  as stated in equation (5.52).

These predictions can be compared with the results of Michell et al. (2008), who observed a dark stripe develop in a region of active aurora. The stripe formed with an initial width of 15 km, then widened to a final width of 55 km. The favoured interpretation for this behaviour is that the dark stripe corresponded to a downward current region (hence low electron precipitation, which caused the dark stripe) and that this broadened due to E-region depletion. The widening rate for this event was reported as  $750 \text{ ms}^{-1}$ . Combination of (5.52) and (5.57) for the broadening speed gives

$$U = \frac{M_P E_y^T}{\sqrt{(1 + \Sigma_P/\Sigma_A)^+ (1 + \Sigma_P/\Sigma_A)^-}}, \quad (6.8)$$

which allows us to check our theory against observations. If  $M_P = 10^4 \text{ m}^2 \text{ s}^{-1} \text{ V}^{-1}$ ,  $\Sigma_P/\Sigma_A \approx 10$  for the undepleted E-region, and  $\Sigma_P/\Sigma_A \approx 1$  in the depleted E-region, then a discontinuity speed of  $750 \text{ ms}^{-1}$  requires a perpendicular electric field in the E-region of approximately  $0.35 \text{ Vm}^{-1}$ . This is a very reasonable value, showing that our expression for the widening speed is consistent with observations. Similarly, Marklund et al. (2001) observed a broadening event with Cluster, with a widening speed in the E-region of approximately  $180 \text{ ms}^{-1}$ : the corresponding perpendicular E-region electric field, approximately  $0.1 \text{ Vm}^{-1}$ , is again very reasonable.

We have also shown that broadening of a large-scale downward current channel naturally produces small-scale Alfvén waves, anchored in the part of the E-region where broadening occurs. Here, broadening collapses the length-scale of the system, forming a discontinuity if the magnetosphere is modelled using ideal MHD. When electron inertial effects are included in the magnetosphere, the discontinuity is smoothed and is trailed by a series of undershoots and overshoots, which evolve as inertial IM-waves. The inertial IM-waves impose electric and magnetic fields at the base of the magnetosphere, driving inertial Alfvén waves that propagate out into the magnetosphere.

In the introduction to this thesis (Section 1.3.4), we quoted two questions that we would try to answer about small-scale Alfvén waves in large scale current systems. The first of these questions was, ‘‘What mechanism generates small-scale waves observed in the magnetosphere?’’ We have now demonstrated that small-scale Alfvén waves, with many of the desired features, are generated by a non-linear interaction between the E-region and overlying magnetosphere during E-region depletion, and we have described this mechanism in detail.

It is not necessary to have a magnetospheric cavity, of the type which leads to ionospheric feedback instability, in order to produce these small-scale waves. If, however, trapping does occur, then it may amplify the waves considerably. Trapping is also likely to cause IM-waves to spread

from their initial location (in the region across which the density cavity broadens), because the inertial Alfvén waves that propagate into the magnetosphere (driven by IM-waves), have a group velocity in the opposite direction to the electric field. Thus, reflected inertial Alfvén waves arrive back at the E-region at a point further into the downward current channel than the point they left from. This behaviour is seen in the results of Streltsov and Lotko (2004) (see Figure 1.8 in this thesis), their simulations including trapping by the IAR.

The second question that we would like to answer is, “What defines the frequency, amplitude and transverse scale-sizes of these waves?” We have shown that small-scale magnetospheric Alfvén waves are driven by inertial IM-waves at the bottom of the magnetosphere. At all times the period of these waves at a fixed location of the E-region is approximately or slightly longer than  $\tau_{IM}$ , the period decreasing asymptotically to  $\tau_{IM}$  as the IM-waves phase-mix to ever shorter length-scales. I expect it to be this period that best characterises small-scale waves produced by broadening of an E-region plasma-density cavity. This period was estimated in Section 5.5.1 as several seconds or tens of seconds, although this may vary considerably, even within the same current system (note the factor seven variation of  $\tilde{\omega}_{IM}$  in our simulations, as shown by Figure 6.8).

We can also predict transverse length-scales of the IM-waves in the E-region. Immediately after they are formed by broadening of an E-region density cavity, inertial IM-waves have a wavelength slightly larger than the electron inertial length at the base of the magnetosphere (likely to be the bottom of the F-region for a stratified model). As time progresses, the inertial IM-waves phase-mix and become increasingly decoupled from one another; we therefore expect the transverse length-scale of these waves be inversely proportional to time (see, e.g., Figure 6.6).

Satellites do not observe IM-waves directly, but rather magnetospheric Alfvén waves driven by IM-waves: IM-waves impose a  $k_y$  and an  $\omega$  at the base of the magnetosphere, which determines  $k_z$  for the Alfvén wave through the inertial Alfvén wave (IAW) dispersion relation (3.34), and fixes the group and phase velocities of the IAW. In a realistic magnetic field geometry, two factors determine how Alfvén wave characteristics map out into the magnetosphere: magnetic field-line geometry and any transverse group velocity caused by inertial terms. The simulation shown in Figure 6.5 illustrates the second of these effects: at the base of the magnetosphere,  $k_y \lambda_e \gtrsim 1$  so electron inertial terms are significant and give the IAW a transverse group velocity in the opposite direction to the phase-velocity; therefore, the IAWs spread out over the downward current channel. In a realistic geomagnetic field, field-lines spread out with increasing distance from the E-region. Hence, the small-scale waves occupy a given fraction of the current channel at altitudes where electron inertial terms are negligible.

There are, therefore, some statements that we can make about satellite observations of small-scale Alfvén waves in large-scale current systems if these waves are produced by ionosphere-

magnetosphere interactions. First, the satellite observations should reveal waves with periods corresponding to the range of  $\tau_{IM}$  present in the E-region footpoint region (allowing for Doppler shift of satellite observations). This is consistent with the results of Karlsson et al. (2004). If multi-point observations are suitably separated in time, it may also be possible to observe the transverse wavelength decreasing in time, as the IM-waves phase-mix. Lastly, we expect that mapping the transverse wavelength of the observed waves to the E-region should yield a length-scale similar to, or less than, the electron inertial length at the base of the F-region.

Our E-region depletion model predicts that small-scale Alfvén waves are more likely to be observed in a large-scale system of field-aligned current that connects with the night-side ionosphere, rather than the day-side ionosphere. Two factors cause this: (i) As discussed in Chapter 4, E-region plasma-density cavities are more likely to form at night than during the day, because the ionisation rate is less at night; it is therefore easier for downward current density to remove electrons from the E-region at an unsustainable rate. Broadening drives the creation of inertial IM-waves, which in turn correspond to small-scale Alfvén waves in the magnetosphere. Therefore, small-scale waves are more likely to be produced at night. (ii) As discussed in Chapter 5, IM-waves are damped by recombination, which means that they have a longer lifetime in the night-side E-region (see Section 5.5.1 for a comparison of estimated damping times).

This conclusion is consistent with observations of small-scale Alfvén waves in large-scale current systems. For example, Karlsson et al. (2004) examined an event that occurred at 20:00 MLT in the Southern Hemisphere during May; for this time of year, this corresponds to night at the ionospheric footpoint. Similarly, the results of Keiling et al. (2005) were made at about 22:00 MLT in the Northern Hemisphere ( $70^\circ$  ILAT) during May; shortly after sunset on the ground.

Despite the success of the E-region depletion model at producing small-scale waves in large-scale current systems, there should be some caution in comparing with existing observations, because of the observed direction of propagation. For example, Karlsson et al. (2004) concluded that the waves they observed were downgoing; Keiling et al. (2005) saw waves in the 9.5–25 s period range (in the spacecraft frame) that were first upgoing and then downgoing. This discrepancy between our model and existing observations could be resolved if small-scale Alfvén waves produced by E-region depletion were reflected somewhere above the spacecraft orbit in both of these cases, or it may be that some as-yet-unknown process was at work on these occasions. Despite this caution, the E-region depletion model remains a powerful means by which the coupled ionosphere-magnetosphere system is sure to produce small-scale Alfvén waves in large-scale systems of field-aligned current.

## Future Work & Conclusion

---

### 7.1 Future Work

#### 7.1.1 Progressing Ionosphere-Magnetosphere Waves

My priority for future work is to press forward with the theory of ionosphere-magnetosphere waves. This theory makes a substantial leap forward in our ability to interpret coupled ionosphere-magnetosphere dynamics, and it is timely, because of growing interest in MI-coupling and advances in numerical simulations (e.g. Streltsov and Lotko (2004, 2008); Streltsov and Pedersen (2010); Sato et al. (2009)) Having established fundamental concepts, such as advection for ideal IM-waves, a large amount of future work is open to us and progress should be rapid. I identify three issues of particular interest for the near future: extension of theory to include a 2D sheet E-region; exploration of the close relationship between IM-waves and ionospheric-feedback instability; and experimental verification of theory using ionospheric heaters.

Extension of IM-wave theory to a 2D sheet E-region (where effects of the Hall current are non-trivial) will be assisted by the existence of suitable models that have already been developed for the closely related study of ionospheric feedback instability. We have, in fact, already recovered the 2D normal-mode IM-wave dispersion relation as a limiting case of the more general IFI dispersion relation stated by Lysak and Song (2002), and examined its properties for an ideal magnetosphere (Section 5.5.3). My suggestion for the future is to attempt to obtain a non-linear ideal advection equation for a 2D sheet E-region, by applying the reflection coefficient approach to the equations of Hasegawa et al. (2010), and confirm (or otherwise) the expectations that we have already drawn from the ideal dispersion relation. This will, however, be complicated by a need to include the fast wave in the magnetosphere. It is also desirable that the strongly inertial limit of the 2D dispersion relation should be closely investigated and characterised.

There is now an appealing opportunity to revise our understanding of ionospheric feedback instability, in the light of IM-wave theory. At present, there is a substantial gap in IFI research, between linear normal-mode theory (Sato, 1978; Lysak and Song, 2002) and more complex numerical simulations (e.g. Lysak and Song (2002); Streltsov and Lotko (2004, 2008); Sato et al. (2009)). A complementary approach that may well help to bridge this gap, is to view arbitrary

E-region perturbations as IM-waves, and explore their evolution when subject to reflection of up-going Alfvén waves at a point somewhere above the E-region. As I commented in Section 5.5.2.1, this novel approach should shed light on advection and growth of disturbances, development of wavetrains and non-linear dynamics. I am also keen to test the conjecture that IFI is produced by resonance between IM-waves and a magnetospheric cavity (by which I mean any trapping region above the E-region, including the IAR).

Lastly, we have the good fortune to be able to verify IM-wave theory using ionospheric heating experiments, and several proposals to test 1D IM-wave theory were made in Section 5.5.4. One issue that was not discussed previously, is how to determine the direction of the background electric field. Traditionally, radar observations have been used to deduce the direction and strength of the electric field, by assuming that motion of ionospheric disturbances is due to  $\mathbf{E} \times \mathbf{B}_0$  drift: clearly this approach is ill-suited to test a theory that disturbances advect with a component in the direction of the electric field. In fact, IM-wave theory suggests that this assumption may need to be revised substantially. Two alternative methods are available: in-situ (rocket) measurements will record the electric field unambiguously, although this adds cost to an experiment; alternatively, the electric field can be assumed from auroral geometry. This second option is perhaps the most appealing, especially since a dark auroral stripe has been identified as offering good conditions to observe IM-waves. Rephrasing the prediction for these experiments, a 1D perturbation in E-region plasma-density with its long axis parallel to the auroral arc, is expected to drift perpendicular to the arc at a speed given by  $v_{IM}$ . (Whilst features will drift in the  $\mathbf{E} \times \mathbf{B}_0$  direction also, invariance in that direction (for a 1D perturbation) will not produce a visible signature.)

### 7.1.2 E-Region Depletion and Small-Scale Waves: Matching Observations

We have had considerable success in developing theory on the formation and widening of E-region density cavities (and simultaneous broadening of downward current channels), and have proposed a detailed E-region depletion model for the origin of small-scale Alfvén waves in large-scale current systems. Advancing these studies further requires a more detailed comparison between theory and observations. Ideally, we should like to have ground and space observations of a broadening event that provide E-region electric fields, E-region plasma-density, and the plasma-density at the base of the F-region (to compute  $v_A$  and  $\lambda_e$ ). Between them, these measurements would provide the theoretical broadening speed, and the range of frequencies and spatial scales likely to be seen in small-scale Alfvén waves. It should also be possible to estimate the decay time for small-scale Alfvén waves. Events can then be compared to theory.

It is also desirable to estimate how commonly small-scale Alfvén waves should appear in observations of large-scale current systems, and in which local-time sectors. This is likely to be a product of the occurrence of downward current densities strong enough to cause broadening,

the lifetime of small-scale Alfvén waves produced by broadening, and the frequency with which satellites pass through large-scale current systems.

### 7.1.3 Amplitude of Small-Scale Alfvén Waves in Large-Scale Current Systems

This thesis has proposed a detailed mechanism that produces small-scale Alfvén waves in large-scale current systems by E-region depletion. We have also successfully identified the factors determining the frequency of these waves ( $\omega_{IM}$ ) and the transverse length-scale ( $\lambda_e$  at the base of the magnetosphere and phase-mixing due to gradients of  $\omega_{IM}$ , combined with propagation of inertial Alfvén waves in the magnetosphere). One question that remains, however, is “What determines the amplitude of these waves?”

A likely candidate is that the amplitude of inertial IM-waves, produced by broadening of an E-region density cavity, increases with the value of  $\beta = \Sigma_{P0}/\Sigma_A$ , the ratio of equilibrium Pedersen conductance to the ideal Alfvén conductance at the base of the magnetosphere. In Section 6.4.2 we observed that depth of the undershoot behind the ‘smoothed discontinuity’ increases with  $\beta$ , probably because of the greater difference between the upper and lower solutions. Unfortunately, when using our inertial code, the simulated ionospheric boundary becomes unstable for large-amplitude IM-waves in the E-region density cavity. It is therefore desirable that some new, more stable method be found to implement this boundary-condition.

A complementary approach, that would not require a new boundary-condition, is to investigate the amplitude of undershoots and overshoots produced by wavebreaking of non-linear wavepackets, using a setup identical to that used in Section 5.4.2. Simulations should quickly reveal how the amplitude of the initial undershoot grows depending on the amplitude of the initial wavepacket, the value of  $\beta$ , and any other parameters we might care to choose. This will inform us about the depletion scenario and could also guide us towards analytic results.

### 7.1.4 Field-Line Resonance

The development that I would most like to arise from my work on field-line resonance is the application of ‘imprinting’ to terrestrial magnetoseismology. The results of Chapter 2 have established that spatial variations of a resonant Alfvén wave, along the resonant surface, can be used to reveal the spatial form of the fast wave that drives the resonance. Initially, I would like to see a feasibility study for this technique that assesses the quality with which variations in ULF wave power can be resolved by ground based instrumentation (I expect that this information is readily available) and investigates methods to map ground observations to the equatorial plane of the magnetosphere (along field-lines). It would be very interesting to see a figure of observed ULF wave energy-

density mapped to the equatorial plane, which could then be compared to the surface plots shown in Figure 2.10. Energy-density figures are likely to vary significantly between events, because of the range of solar-wind drivers thought to excite ULF pulsations: for example Kelvin-Helmholtz instability at the magnetopause (Mann et al., 1999), broadband drivers that preferentially excite fast wave eigenmodes (Kivelson and Southwood, 1985), or solar wind perturbations running along the flank of the magnetosphere (Wright and Rickard, 1995). It is quite possible that each driver will have a different appearance in fast wave energy-density, and this can be compared with other seismological signatures (e.g. Mann and Wright (1999)). For events with several distinct resonances, we may also wish to identify the spatial form of the fast wave driving each resonance, perhaps by examining energy-density filtered into appropriate frequency bands.

The present work has also prepared the way for a full solution to the 3D hydromagnetic box model of field-line resonance (straight uniform equilibrium magnetic field, with 3D variation of equilibrium Alfvén speed through the density profile). This problem is a keystone of field-line resonance theory, and, as such, would be a worthwhile project as well as an interesting exercise in mathematics. Schulze-Berge et al. (1992) provide a framework for the solution, density variation along field-lines can be handled by the methods of Thompson and Wright (1993) and two dimensional variation of field-line Alfvén eigenfrequencies can be treated using the methods of this thesis. Despite the appeal of this problem, it is anticipated that the key features of this ‘complete’ solution have already been obtained by splitting it into its underlying parts (variation along the magnetic field considered by Thompson and Wright (1993), and variation across the field considered in this thesis and Russell and Wright (2010)).

## 7.2 Conclusion

The core message of this thesis is that interactions between parts of a system, such as the solar wind, magnetosphere and ionosphere, are often richer and more fascinating than the workings of those parts in isolation. Whether we consider field-line resonance, or magnetosphere-ionosphere coupling, the combination of simpler parts has repeatedly produced surprises about our planet and its space-environment. It is, to my mind, a triumph to have opened more areas for further study than we have successfully concluded, and, in this regard, it has been difficult to write this thesis because it has meant pausing my own research. I look forward to taking this up again: in my opinion, we who research these topics can look forward to a busy and exciting future.

---

## Bibliography

---

- A. Aikio, K. Mursula, S. Buchert, F. Forme, O. Amm, G. Marklund, M. Dunlop, D. Fontaine, A. Vaivads, and A. Fazakerley. Temporal evolution of two auroral arcs as measured by the Cluster satellite and coordinated ground-based instruments. *Annales Geophysicae*, 22:4089–4101, December 2004.
- A. T. Aikio, T. Lakkala, A. Kozlovsky, and P. J. S. Williams. Electric fields and currents of stable drifting auroral arcs in the evening sector. *Journal of Geophysical Research (Space Physics)*, 107:1424, December 2002. doi: 10.1029/2001JA009172.
- H. Alfvén. Existence of Electromagnetic-Hydrodynamic Waves. *Nature*, 150:405–406, October 1942. doi: 10.1038/150405d0.
- W. Allan, S. P. White, and E. M. Poulter. Impulse-excited hydromagnetic cavity and field-line resonances in the magnetosphere. *Plan. Space Sci.*, 34:371–385, April 1986. doi: 10.1016/0032-0633(86)90144-3.
- B. J. Anderson, M. J. Engebretson, S. P. Rounds, L. J. Zanetti, and T. A. Potemra. A statistical study of Pc 3-5 pulsations observed by the AMPTE/CCE magnetic fields experiment. I - Occurrence distributions. *J. Geophys. Res.*, 95:10495–10523, July 1990. doi: 10.1029/JA095iA07p10495.
- M. J. Aschwanden. Three Criteria to Discriminate between Elementary and Composite Coronal Loops. *Astrophys. J. Letts.*, 634:L193–L196, December 2005. doi: 10.1086/499041.
- M. J. Aschwanden and R. W. Nightingale. Elementary Loop Structures in the Solar Corona Analyzed from TRACE Triple-Filter Images. *Astrophys. J.*, 633:499–517, November 2005. doi: 10.1086/452630.
- M. J. Aschwanden, L. Fletcher, C. J. Schrijver, and D. Alexander. Coronal Loop Oscillations Observed with the Transition Region and Coronal Explorer. *Astrophys. J.*, 520:880–894, August 1999. doi: 10.1086/307502.
- G. Atkinson. Auroral arcs: Result of the interaction of a dynamic magnetosphere with the ionosphere. *J. Geophys. Res.*, 75:4746–4755, 1970. doi: 10.1029/JA075i025p04746.
- C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. Advanced Mathematical Methods for Scientists and Engineers, New York: McGraw-Hill, 1978, 1978.
- E. M. Blixt and A. Brekke. A model of currents and electric fields in a discrete auroral arc. *Geophys. Res. Lett.*, 23:2553–2556, 1996. doi: 10.1029/96GL02378.

- T. J. M. Boyd and J. J. Sanderson. *Plasma dynamics*. 1969.
- R. C. Carrington. Description of a Singular Appearance seen in the Sun on September 1, 1859. *MNRAS*, 20:13–15, November 1859.
- L. Chen and A. Hasegawa. Plasma heating by spatial resonance of Alfvén wave. *Physics of Fluids*, 17:1399–1403, July 1974. doi: 10.1063/1.1694904.
- A. P. Cran-McGreehin, A. N. Wright, and A. W. Hood. Ionospheric depletion in auroral downward currents. *Journal of Geophysical Research (Space Physics)*, 112(A11):10309, October 2007. doi: 10.1029/2007JA012350.
- A. de Groof, W. J. Tirry, and M. Goossens. Random driven fast waves in coronal loops. I. Without coupling to Alfvén waves. *Astron. Astrophys.*, 335:329–340, July 1998.
- R. A. Doe, J. F. Vickrey, and M. Mendillo. Electrodynamical model for the formation of auroral ionospheric cavities. *J. Geophys. Res.*, 100:9683–9696, June 1995. doi: 10.1029/95JA00001.
- J. W. Dungey. Electrodynamics of the Outer Atmosphere. In *Physics of the Ionosphere*, page 229, 1955.
- P. M. Edwin and B. Roberts. Wave propagation in a magnetic cylinder. *Solar Phys.*, 88:179–191, October 1983. doi: 10.1007/BF00196186.
- B. J. Fraser. ULF Waves — A Historical Note. In K. Takahashi, P. J. Chi, R. E. Denton, and R. L. Lysak, editors, *Magnetospheric ULF Waves: Synthesis and New Directions*, pages 5–10. American Geophysical Union, 2006.
- H. Hasegawa, N. Ohno, and T. Sato. Simulation of feedback instability in the coupled magnetosphere-ionosphere system. *Journal of Geophysical Research (Space Physics)*, 115: 8304, August 2010. doi: 10.1029/2009JA015093.
- R. Hodgson. On a curious Appearance seen in the Sun. *MNRAS*, 20:15–16, November 1859.
- T. E. Holzer and T. Sato. Quiet auroral arcs and electrodynamic coupling between the ionosphere and the magnetosphere. 2. *J. Geophys. Res.*, 78:7330–7339, 1973. doi: 10.1029/JA078i031p07330.
- A. W. Hood. Oscillatory structure in viscous shocks. private communication, 2010.
- U. S. Inan, H. C. Chang, R. A. Helliwell, W. L. Imhof, J. B. Reagan, and M. Walt. Precipitation of radiation belt electrons by man-made waves A comparison between theory and measurement. *J. Geophys. Res.*, 90:359–369, January 1985. doi: 10.1029/JA090iA01p00359.
- T. Johansson, S. Figueiredo, T. Karlsson, G. Marklund, A. Fazakerley, S. Buchert, P. Lindqvist, and H. Nilsson. Intense high-altitude auroral electric fields - temporal and spatial characteristics. *Annales Geophysicae*, 22:2485–2495, July 2004.
- T. Karlsson and G. T. Marklund. Simulations of effects of small-scale auroral current closure in the return current region. *Phys. Space Plasmas*, 15:401, 1998.
- T. Karlsson and G. T. Marklund. Results from the DC electric field experiment on the Freja satellite. *Advances in Space Research*, 23:1657–1665, 1999. doi: 10.1016/S0273-1177(99)00375-0.
- T. Karlsson, G. Marklund, S. Figueiredo, T. Johansson, and S. Buchert. Separating spatial and temporal variations in auroral electric and magnetic fields by Cluster multipoint measurements.

- Annales Geophysicae*, 22:2463–2472, July 2004.
- A. Keiling, G. K. Parks, J. R. Wygant, J. Dombek, F. S. Mozer, C. T. Russell, A. V. Streltsov, and W. Lotko. Some properties of Alfvén waves: Observations in the tail lobes and the plasma sheet boundary layer. *Journal of Geophysical Research (Space Physics)*, 110(A9):10, October 2005. doi: 10.1029/2004JA010907.
- M. G. Kivelson and D. J. Southwood. Resonant ULF waves - A new interpretation. *Geophys. Res. Lett.*, 12:49–52, January 1985. doi: 10.1029/GL012i001p00049.
- D.-H. Lee and R. L. Lysak. Magnetospheric ULF wave coupling in the dipole model - The impulsive excitation. *J. Geophys. Res.*, 94:17097–17103, December 1989. doi: 10.1029/JA094iA12p17097.
- D.-H. Lee, R. L. Lysak, and Y. Song. Field line resonances in a nonaxisymmetric magnetic field. *J. Geophys. Res.*, 105:10703–10712, May 2000. doi: 10.1029/1999JA000295.
- R. L. Lysak. Feedback instability of the ionospheric resonant cavity. *J. Geophys. Res.*, 96:1553–1568, February 1991. doi: 10.1029/90JA02154.
- R. L. Lysak and Y. Song. Energetics of the ionospheric feedback interaction. *Journal of Geophysical Research (Space Physics)*, 107:1160, August 2002. doi: 10.1029/2001JA000308.
- R. L. Lysak and Y. Song. Propagation of kinetic Alfvén waves in the ionospheric Alfvén resonator in the presence of density cavities. *Geophys. Res. Lett.*, 35:20101, October 2008. doi: 10.1029/2008GL035728.
- I. R. Mann and A. N. Wright. Diagnosing the excitation mechanisms of Pc5 magnetospheric flank waveguide modes and FLRs. *Geophys. Res. Lett.*, 26:2609–2612, August 1999. doi: 10.1029/1999GL900573.
- I. R. Mann, A. N. Wright, and P. S. Cally. Coupling of magnetospheric cavity modes to field line resonances: A study of resonance widths. *J. Geophys. Res.*, 100:19441–19456, October 1995. doi: 10.1029/95JA00820.
- I. R. Mann, A. N. Wright, K. J. Mills, and V. M. Nakariakov. Excitation of magnetospheric waveguide modes by magnetosheath flows. *J. Geophys. Res.*, 104:333–354, January 1999. doi: 10.1029/1998JA900026.
- G. T. Marklund, N. Ivchenko, T. Karlsson, A. Fazakerley, M. Dunlop, P.-A. Lindqvist, S. Buchert, C. Owen, M. Taylor, A. Vaivalds, P. Carter, M. André, and A. Balogh. Temporal evolution of the electric field accelerating electrons away from the auroral ionosphere. *Nature*, 414:724–727, December 2001.
- P. C. H. Martens, J. W. Cirtain, and J. T. Schmelz. The Inadequacy of Temperature Measurements in the Solar Corona through Narrowband Filter and Line Ratios. *Astrophys. J. Letts.*, 577: L115–L117, October 2002. doi: 10.1086/344254.
- R. G. Michell, K. A. Lynch, and H. C. Stenbaek-Nielsen. Ground-based observational signature of a downward current channel in an active auroral arc. *Geophys. Res. Lett.*, 35:8101, April 2008. doi: 10.1029/2008GL033596.
- T. Nagata, S. Kobun, and T. Iijima. Geomagnetically Conjugate Relationships of Giant Pulsa-

- tions at Syowa Base, Antarctica, and Reykjavik, Iceland. *J. Geophys. Res.*, 68:4621, August 1963.
- V. M. Nakariakov, L. Ofman, E. E. Deluca, B. Roberts, and J. M. Davila. TRACE observation of damped coronal loop oscillations: Implications for coronal heating. *Science*, 285:862–864, August 1999. doi: 10.1126/science.285.5429.862.
- New York Times. The Auroral Display in Boston, Sept, 3 1859a.
- New York Times. Auroral Phenomena, Sept, 5 1859b.
- L. Ofman, M. Balikhin, C. T. Russell, and M. Gedalin. Collisionless relaxation of ion distributions downstream of laminar quasi-perpendicular shocks. *Journal of Geophysical Research (Space Physics)*, 114(A13):9106, September 2009. doi: 10.1029/2009JA014365.
- G. Paschmann, S. Haaland, and R. Treumann. Auroral Plasma Physics. *Space Science Reviews*, 103, April 2002. doi: 10.1023/A:1023088315789.
- V. L. Patel. Low frequency hydromagnetic waves in the magnetosphere: Explorer XII. *Plan. Space Sci.*, 13:485, June 1965. doi: 10.1016/0032-0633(65)90163-7.
- H. R. Radoski. A note on the problem of hydromagnetic resonances in the magnetosphere. *Plan. Space Sci.*, 19:1012–1013, August 1971. doi: 10.1016/0032-0633(71)90152-8.
- V. Rezniaia and J. C. Samson. Quasi-periodic oscillations: Resonant shear Alfvén waves in neutron star magnetospheres. *Astron. Astrophys.*, 436:999–1008, June 2005. doi: 10.1051/0004-6361:20041796.
- G. J. Rickard and A. N. Wright. Alfvén resonance excitation and fast wave propagation in magnetospheric waveguides. *J. Geophys. Res.*, 99:13455–13464, January 1994. doi: 10.1029/94JA00674.
- B. Roberts. Magnetohydrodynamic Waves. In E. R. Priest, editor, *Solar System Magnetic Fields*, pages 37–79. D. Reidel Publishing Company, 1985.
- T. R. Robinson, R. Strangeway, D. M. Wright, J. A. Davies, R. B. Horne, T. K. Yeoman, A. J. Stocker, M. Lester, M. T. Rietveld, I. R. Mann, C. W. Carlson, and J. P. McFadden. FAST observations of ULF waves injected into the magnetosphere by means of modulated RF heating of the auroral electrojet. *Geophys. Res. Lett.*, 27:3165–3168, October 2000. doi: 10.1029/2000GL011882.
- M. S. Ruderman. The resonant damping of oscillations of coronal loops with elliptic cross-sections. *Astron. Astrophys.*, 409:287–297, October 2003. doi: 10.1051/0004-6361:20031079.
- M. S. Ruderman and B. Roberts. The Damping of Coronal Loop Oscillations. *Astrophys. J.*, 577: 475–486, September 2002. doi: 10.1086/342130.
- A. J. B. Russell and A. N. Wright. Resonant absorption with 2D variation of field line eigenfrequencies. *Astron. Astrophys.*, 511:A17, February 2010. doi: 10.1051/0004-6361/200912669.
- A. J. B. Russell, A. N. Wright, and A. W. Hood. Self-consistent ionospheric plasma density modifications by field-aligned currents: Steady state solutions. *Journal of Geophysical Research (Space Physics)*, 115(A14):4216, April 2010. doi: 10.1029/2009JA014836.
- J. C. Samson, J. A. Jacobs, and G. Rostoker. Latitude-Dependent Characteristics of Long-

- Period Geomagnetic Micropulsations. *J. Geophys. Res.*, 76:3675–3683, 1971. doi: 10.1029/JA076i016p03675.
- J. C. Samson, B. G. Harrold, J. M. Ruohoniemi, R. A. Greenwald, and A. D. M. Walker. Field line resonances associated with MHD waveguides in the magnetosphere. *Geophys. Res. Lett.*, 19:441–444, March 1992. doi: 10.1029/92GL00116.
- T. Sato. A theory of quiet auroral arcs. *J. Geophys. Res.*, 83:1042–1048, March 1978. doi: 10.1029/JA083iA03p01042.
- T. Sato and T. E. Holzer. Quiet auroral arcs and electrodynamic coupling between the ionosphere and the magnetosphere. 1. *J. Geophys. Res.*, 78:7314–7329, 1973. doi: 10.1029/JA078i031p07314.
- T. Sato, H. Hasegawa, and N. Ohno. Macro-micro interlocked simulation algorithm: an exemplification for aurora arc evolution. *Computational Science and Discovery*, 2(1):015007, January 2009. doi: 10.1088/1749-4699/2/1/015007.
- J. T. Schmelz. Are Coronal Loops Isothermal? *Astrophys. J. Letts.*, 578:L161–L164, October 2002. doi: 10.1086/344715.
- J. T. Schmelz, R. T. Scopes, J. W. Cirtain, H. D. Winter, and J. D. Allen. Observational Constraints on Coronal Heating Models Using Coronal Diagnostics Spectrometer and Soft X-Ray Telescope Data. *Astrophys. J.*, 556:896–904, August 2001. doi: 10.1086/321588.
- J. T. Schmelz, J. E. Beene, K. Nasraoui, H. T. Blevins, P. C. H. Martens, and J. W. Cirtain. The Effect of Background Subtraction on the Temperature of EIT Coronal Loops. *Astrophys. J.*, 599:604–614, December 2003. doi: 10.1086/379212.
- J. T. Schmelz, K. Nasraoui, V. L. Richardson, P. J. Hubbard, C. R. Nevels, and J. E. Beene. All Coronal Loops Are the Same: Evidence to the Contrary. *Astrophys. J. Letts.*, 627:L81–L84, July 2005. doi: 10.1086/431950.
- S. Schulze-Berge, S. Gowley, and L. Chen. Theory of field line resonances of standing shear Alfvén waves in three-dimensional inhomogeneous plasmas. *J. Geophys. Res.*, 97:3219–3222, March 1992. doi: 10.1029/91JA02804.
- H. J. Singer, W. J. Hughes, and C. T. Russell. Standing hydromagnetic waves observed by ISEE 1 and 2 - Radial extent and harmonic. *J. Geophys. Res.*, 87:3519–3529, May 1982. doi: 10.1029/JA087iA05p03519.
- D. J. Southwood. Some features of field line resonances in the magnetosphere. *Plan. Space Sci.*, 22:483–491, March 1974. doi: 10.1016/0032-0633(74)90078-6.
- B. Stewart. On the Great Magnetic Disturbance Which Extended from August 28 to September 7, 1859, as Recorded by Photography at the Kew Observatory. *Royal Society of London Philosophical Transactions Series I*, 151:423–430, 1861.
- A. V. Streltsov. Proposal: Multi-scale electrodynamics of magnetosphere-ionosphere interactions at high latitudes. research proposal, 2008.
- A. V. Streltsov and T. Karlsson. Small-scale, localized electromagnetic waves observed by Cluster: Result of magnetosphere-ionosphere interactions. *Geophys. Res. Lett.*, 35:22107, November

2008. doi: 10.1029/2008GL035956.
- A. V. Streltsov and W. Lotko. Multiscale electrodynamics of the ionosphere-magnetosphere system. *Journal of Geophysical Research (Space Physics)*, 109(A18):9214, September 2004. doi: 10.1029/2004JA010457.
- A. V. Streltsov and W. Lotko. Ultra-low-frequency electrodynamics of the magnetosphere-ionosphere interaction. *Journal of Geophysical Research (Space Physics)*, 110(A9):8203, August 2005. doi: 10.1029/2004JA010764.
- A. V. Streltsov and W. Lotko. Coupling between density structures, electromagnetic waves and ionospheric feedback in the auroral zone. *Journal of Geophysical Research (Space Physics)*, 113(A12):5212, May 2008. doi: 10.1029/2007JA012594.
- A. V. Streltsov and T. R. Pedersen. An alternative method for generation of ULF waves by ionospheric heating. *Geophys. Res. Lett.*, 37:14101, July 2010. doi: 10.1029/2010GL043543.
- A. V. Streltsov, T. R. Pedersen, E. V. Mishin, and A. L. Snyder. Ionospheric feedback instability and substorm development. *Journal of Geophysical Research (Space Physics)*, 115:7205, July 2010. doi: 10.1029/2009JA014961.
- M. Sugiura and C. R. Wilson. Oscillation of the Geomagnetic Field Lines and Associated Magnetic Perturbations at Conjugate Points. *J. Geophys. Res.*, 69:1211–1216, April 1964. doi: 10.1029/JZ069i007p01211.
- K. Takahashi and B. J. Anderson. Distribution of ULF energy ( $f$  is less than 80 mHz) in the inner magnetosphere - A statistical analysis of AMPTE CCE magnetic field data. *J. Geophys. Res.*, 97:10751, July 1992. doi: 10.1029/92JA00328.
- K. Takahashi, B. J. Anderson, and S.-i. Ohtani. Multisatellite study of nightside transient toroidal waves. *J. Geophys. Res.*, 101:24815–24826, November 1996. doi: 10.1029/96JA02045.
- J. Terradas, R. Oliver, and J. L. Ballester. Damped Coronal Loop Oscillations: Time-dependent Results. *Astrophys. J.*, 642:533–540, May 2006. doi: 10.1086/500730.
- J. Terradas, I. Arregui, R. Oliver, J. L. Ballester, J. Andries, and M. Goossens. Resonant Absorption in Complicated Plasma Configurations: Applications to Multistranded Coronal Loop Oscillations. *Astrophys. J.*, 679:1611–1620, June 2008. doi: 10.1086/586733.
- M. J. Thompson and A. N. Wright. Resonant Alfvén wave excitation in two-dimensional systems - Singularities in partial differential equations. *J. Geophys. Res.*, 98:15541–15551, September 1993. doi: 10.1029/93JA00791.
- C. Q. Wei, J. C. Samson, R. Rankin, and P. Frycz. Electron inertial effects on geomagnetic field line resonances. *J. Geophys. Res.*, 99:11265–11276, June 1994. doi: 10.1029/94JA00582.
- A. N. Wright and G. J. Rickard. ULF pulsations driven by magnetopause motions: Azimuthal phase characteristics. *J. Geophys. Res.*, 100:23703–23710, December 1995. doi: 10.1029/95JA01765.